Acquisition of Procedural Skills in Preregistration Physiotherapy Education Comparing Mental Practice Against No Mental Practice: The Learning of Procedures in Physiotherapy Education Trial – A Development of Concept Study

Journal of Medical Education and Curricular Development
Volume 7: 1–13
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2382120520927382
SAGE

Karl Martin Sattelmayer[1,2], Kavi C. Jagadamma[1], Roger Hilfiker[2] and Gillian Baer[1]

[1]Physiotherapy, School of Health Sciences, Queen Margaret University, Edinburgh, UK. [2]School of Health Sciences, University of Applied Sciences and Arts Western Switzerland Valais (HES-SO Valais-Wallis), Leukerbad, Switzerland.

ABSTRACT

INTRODUCTION: Procedural skills are a central element in the education of physiotherapists. Procedural skills relate to the execution of a practical task. An educational intervention, which can be used to support skill acquisition of procedural skills, is mental practice (MP). Several studies have investigated the use of MP or imaging in medical education. This pilot study evaluated the application of MP on the acquisition of procedural skills in physiotherapy education.

METHODS: This pilot randomised controlled study recruited a convenience sample of 37 BSc physiotherapy student participants. Two different complex task procedures (transfer and vestibular rehabilitation) were trained during this study. Participants in both the transfer (task procedure 1) and the vestibular rehabilitation (task procedure 2) arm of the study were randomly assigned to either MP or no MP.

RESULTS: For the transfer task, median performance at post-acquisition testing showed a moderate effect size in favour of the group using MP (*r*: −0.3), but the findings were not statistically significant (*P*: 0.2). Similar results were found for the vestibular rehabilitation task (*r*: 0.29; *P*: 0.21). In addition, the self-reported confidence was higher in the MP group.

CONCLUSION: Moderate effect sizes were identified in favour of MP at post-acquisition testing. In addition, the between-group difference was higher than the minimally important difference. The feasibility of the study was high based on quantitative feasibility measures such as the recruitment rate. Both these findings suggest larger well-powered studies should be considered to confirm the findings of this pilot study.

KEYWORDS: Learning [F02.463.425], motor skills [F02.808.260], psychomotor performance [F02.808], public health professional education [I02.358.556], mental practice

RECEIVED: November 25, 2019. **ACCEPTED:** April 21, 2020.

TYPE: Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Karl Martin Sattelmayer, School of Health Sciences, University of Applied Sciences and Arts Western Switzerland Valais (HES-SO Valais-Wallis), 3954 Leukerbad, Switzerland. Email: martin.sattelmayer@hevs.ch

Introduction

The education of procedural skills is an essential element in the training of future physiotherapists.[1] Procedural skills relate to the execution of a practical task (eg, a manual mobilisation technique). Procedures can be classified into different categories, such as procedures related to diagnostic or interventional processes. Incorrectly performed procedures can compromise the effectiveness of therapeutic interventions or may lead to unwanted problems and side effects for patients and physiotherapists. For instance, Gorrell et al[2] reported serious adverse events after spinal manipulation in two studies and minor adverse events in 61 studies (in total, 368 studies were included in their systematic review).

The importance of correctly acquiring procedural skills in health professionals education can be seen by the increasing volume of published studies on the topic. New procedures in health care are developed constantly, which requires that educational programmes either increase the amount of taught procedures in their curricula or select new procedures to teach and discard existing procedures.[3] This dilemma highlights the need for effective and feasible methods to support learners and educators.[4] In addition, it has been reported that the ability to perform procedures such as a physical examination has deteriorated over the last decade.[5] Several methods have been introduced to respond to this challenge. For example, Internet-based learning applications[6] or virtual reality simulation[7] are increasingly used in health professions education. However, the use of those technology-based methods requires considerable resources.[8] An educational intervention requiring less resources and feasible for common use in an educational setting is mental practice (MP).[9]

Mental practice was defined by Schmidt and Lee[10] as 'the performance of a task is mentally rehearsed in the absence of overt physical practice'. Mental practice is a relatively broad concept and can include techniques such as thinking about a

motor skill, and it can also involve imagery techniques (kinaesthetic or visual or imagery).

Traditionally, MP has been used in numerous sport disciplines to enhance the acquisition of complex motor skills.[11] In a frequently quoted systematic review, Landers[12] found that MP had a considerable effect size of 0.48 on motor skill acquisition. More than 60 studies were analysed, and a broad range of motor skills, such as juggling and dart throwing, were included. In addition, Feltz et al[13] reported in a follow-up meta-analysis that MP can be used to accelerate the process of motor skill acquisition. More recent findings by Wohldmann et al[14] showed that MP can be at least as effective as physical practice for motor skill acquisition in sport.

Most of the published findings on the effectiveness of MP are based on studies that report movement skills in sports disciplines. It is, therefore, important to evaluate the existing findings of MP in the training of tasks performed by health professionals, which can be more variable than a sport movement skill. A systematic review published by Sattelmayer et al[15] presented evidence that MP can be applied effectively in medical education. Eight randomised controlled trials were included in the analysis, all reporting on procedures in medical education such as laparoscopic surgery or cricothyrotomy. A moderate effect size (standardised mean difference [SMD]: 0.43) was found in favour of MP on post-acquisition tests. A small effect (statistically not significant) was identified in favour of MP on a retention test (SMD: 0.2). However, no study was identified reporting on the application of MP for skill acquisition in physiotherapy education.

Study Aim and Objectives

The aim of this pilot study was to evaluate the application of MP on the acquisition of procedural skills in physiotherapy education. Several parameters required investigation within this study.

The primary objective was to estimate the effectiveness of MP on performance measures of procedural skills. The secondary objective was to analyse the feasibility of this study with several feasibility criteria.

Methods

Design

A 'development of concept'[16] pilot randomised controlled trial design was used. This design implies the use of a control group with a (active) comparator, randomisation procedures, and masked outcomes.[16] All these elements are relevant to adequately estimate the effectiveness of MP. In addition, feasibility issues can be explored which are relevant for the decision whether one should go on with larger trials (eg, was it feasible to recruit sufficient participants?).

Participants

A convenience sample of 37 year 2 and year 3 BSc physiotherapy student participants from the University of Applied Sciences Western Switzerland Valais (UAS Valais) were invited to participate in the Learning of Procedures in Physiotherapy Education (LEArN) trial (ie, for task procedure 1 and 2). A research assistant, not involved in teaching the procedures, approached the potential participants who had 4 days to consider participation. Participants provided written informed consent.

Criteria for considering participants for this study

All second and third year undergraduate physiotherapy students at the UAS Valais were considered eligible for inclusion. Prior formal training in the procedures (transfer and vestibular rehabilitation) was used as exclusion criterion.

Ethics

The study received approval from the Commission cantonale d'éthique de la recherche sur l'être humain (CER-VD) Switzerland (December 8, 2016) and the Research Ethics committee of Queen Margaret University (February 17, 2017).

Randomisation

Participants in both the transfer (task procedure 1) and the vestibular rehabilitation (task procedure 2) arm of the study were randomly assigned, via a computer-generated random number table, to either MP (group A) or no mental practice (nMP; group B). Previous performance during university-based procedural skills examinations was regarded as a potential confounding variable, and therefore, stratification was used to ensure an equal distribution of this variable. Four strata, based on previous performance, were generated (ie, high, above average, below average, and low performer). The randomisation was performed in R using the blockrand package[17] by an independent person not employed at the UAS Valais. The block size was set to 4.

Allocation concealment. Details of the participants were provided to the person performing the randomisation by email, ie, the allocation sequence was concealed by a central allocation service. The allocation sequence was concealed to the personnel performing the intervention until participants were officially registered to a study arm.

Intervention

Task procedures used. Two different complex procedures were trained during this study. The procedure for the participants in the third year of their training was a transfer to the ground for a person with a stroke (task procedure 1). The participants learned this procedure for the first time but had prior training of a similar task procedure (transfer in a sitting position). The procedure for the participants in the second year of their training was a set of procedures from vestibular rehabilitation (task procedure 2). Participants were novices for this task

procedure. They had no training in this or a similar procedure before participating in the study and can therefore be classified as novices.

Both procedures are complex for novice practitioners and consist of multiple movement parts that require a set of specific movement skills. The reason for conducting this study with two different procedures was that both procedures vary to some degree (eg, regarding their movement parts, complexity, or previous experience of the students). Therefore, it was anticipated that this might provide more insight into the application and analysis of motor learning principles.

Training of procedures. The procedural skills training was undertaken by one educator and lasted for 1.5 to 2 hours. First, a general introduction to the procedure was provided, which covered theory and a general guideline. Then, the specific procedure (dependent on group allocation) was demonstrated. Group size for the procedural skills training was 10 participants or less.

- Group A received a procedural skills training which consisted of MP and physical practice (group MP).
- Group B received a procedural skills training which consisted of physical practice and nMP (group nMP)

The training of group B consisted of physical training. After the instruction and demonstration of the procedure, participants were asked to practice in pairs. One participant was asked to perform the procedure as a physiotherapist and one participant was asked to perform the procedure as a patient. After one practice trial, the participants switched their roles and performed a second practice trial. This was continued until the end of the practice session. The training of group A was also performed in pairs. After each practice trial, participants switched their roles. In addition to the physical practice, they were asked to perform an MP trial as well (ie, they mentally rehearsed the procedure in the absence of movement). The MP consisted of the following elements: (1) the training was conducted in a group with group supervision, (2) the timing of the MP was before physical practice (ie, one practice trial consisted of a mental procedure execution followed by a physical procedure execution [after this, a new practice trial started]), (3) the main focus of the MP script was on motor tasks (eg, performing a weight shift or controlling the head position during a specific procedure part), (4) cues from different modalities were included into the MP scripts; kinaesthetic cues (ie, cues related to moving and feeling oneself or the patient during the procedure), collaborative cues (ie, procedure elements that should be performed in collaboration between physiotherapist and patients), and cognitive cues (ie, cues related to thinking and decision-making processes that occur during the procedure) were integrated into the MP script, and (5) an internal, first-person perspective was instructed for the MP. The MP script for the transfer procedure was developed and validated a priori using an approach presented by Arora et al.[18] That is, protocol analysis and expert thought were used to develop an MP script.[19] Three experienced physiotherapists were interviewed and asked to perform a cognitive walk-through of the procedure. The interviews were analysed for relevant cues. Based on these, a preliminary MP script was developed. In the second phase, the MP script was validated using a pre-post measurement study design. Within this phase, 11 participants rated their ability to image the procedure on a movement imagery questionnaire.[20] This was followed by a single MP training session using the MP script, followed by a second movement imagery test. The analysis showed that the ability to image the procedure increased considerably after the MP training.[21] The MP script for task procedure 2 was developed using a similar approach, but the script was not validated in a pilot study. To clarify the MP intervention, an overview of the transfer script is provided in Appendix 1. The duration of the training in both groups was set to 1.5 hours but could be extended to 2 hours in case of open questions or if participants wanted to complete their practice trial. The training was led in both groups by the same educator (K.M.S.), who has several years of teaching experience with the task procedures. No other educator participated in this study or was part of the research team. The MP training was prepared according to the principles stated above and trained over several practice sessions with a total duration of 20 hours and faculty members experienced with the use of MP provided feedback.

Outcomes. The outcome measurement was performed by rating video recordings of participants performing the trained procedure.

Endpoints. Two different endpoints, to assess learning, were used. A post-acquisition test on a peer participant was conducted immediately after the procedural skills training (T1). That is, the respective task procedure was performed on the respective peer participant and the performance was video recorded. The same procedure was performed 2 weeks after the acquisition phase to ascertain retention (T2). At T2, the procedure was performed on a trained standardised 'patient' (ie, healthy volunteers) simulating a person suffering from stroke or a vestibular disorder. Procedures at T1 and T2 were video recorded and performance was evaluated based on predetermined criteria.

Outcome measures. Performance of the procedure and, therefore, an indication of the effectiveness of the educational interventions were measured with the outcome measures given below at endpoints T1 and T2.

Procedural skills. The primary outcome for effectiveness was procedural skills. The construct procedural skills was measured with the Assessment for Procedural Skills in Physiotherapy Education with 29 items (APSPT 29), which is a measurement instrument designed to evaluate the performance of procedures in physiotherapy education. The APSPT 29 is an ordinal scaled generic assessment for procedural skills (ie, it can

be used for various procedures in physiotherapy) and contains 6 subcategories (preparation, knowledge and decision-making, communication, safety, procedure execution, and comfort) with a total of 29 items. The APSPT 29 has been investigated in a previous pilot study. Preliminary Rasch analysis showed that all 29 items had adequate fit to the construct 'procedural skills' regarding infit and outfit mean square statistics. Furthermore, the internal consistency of the APSPT was analysed with a Cronbach α of 0.95.[22] The APSPT 29 total score can range between 0 and 116 points.

Procedure-specific aspects. A procedure-specific checklist (PSC) was used to evaluate specific aspects of the procedures. The checklist contained critical points of the procedures. More specifically, procedure-specific elements were rated on a trichotomous scale. The PSC total score could range between 0 and 24 points (transfer procedure) and between 0 and 26 points for the vestibular rehabilitation procedure. Higher scores indicated a better performance. The PSC was developed based on the content of the procedural skills and was not validated a priori.

Response time. Response time is a core measure of procedural skills with important educational implications. Response time has been analysed to be a sensitive measure of change, which can detect training-induced changes in novices and experts in motor learning research.[23] Furthermore, many studies evaluating motor learning principles in health professions education used response time as an outcome.[22] The response time was the time that a participant needed to perform the procedure measured in seconds.

Self-reported confidence. Three items from a 6-item questionnaire (surgical skills attitude questionnaire) designed to evaluate self-reported confidence of procedural skills in medical education[24] were used. Edwards et al[25] reported an adequate reliability of the questionnaire and a positive correlation with student performance. The omitted items measured confidence in surgical procedures and, therefore, were not relevant to physiotherapy education. The total score could range between 0 and 12 points.

Blinding of outcome assessors. An independent physiotherapist with experience in the procedures (not employed at the UAS Valais) evaluated the performances of the tasks on the video recordings. The following outcomes were completed by the blinded assessor: procedural skills, procedure-specific aspects, response time, and failure rate. The physiotherapist was unaware of the group allocation of the participants. The remaining outcomes were evaluated by the first author.

Feasibility criteria. The feasibility of the MP intervention was evaluated with several parameters. The criteria were chosen to provide information about the feasibility of methods and procedures for later use on a larger scale. Feasibility criteria for this pilot study were identified in Thabane et al[26] and Van Teijlingen and Hundley.[27]

Recruitment rate. The recruitment rate was calculated by subtracting the recruited participants from the participants meeting the eligibility criteria. A recruitment rate lower than 50% would raise doubts about the feasibility of a larger study.

Failure rate. A failure to perform the procedure was investigated for all participants and for each specific group. Previous experience led to the expectation that the majority of the participants should have been able to perform the procedure adequately after procedural skills training. Feasibility of the approach to acquiring the procedural skill might be considered questionable if more than 40% of the participants could not perform the procedure effectively. The independent video rater classified each performance as adequate or not adequate based on a general impression of the performance. This evaluation was a subjective overall assessment of the performance. Ten Cate and Regehr[28] argue that expert judgement should be recognised as a central part of the assessment in health professions education and can provide valuable information in addition to more objective outcome measures.

Feasibility of the procedural skills training session. Following the training, the educator completed a short questionnaire regarding the feasibility of the training sessions (eg, difficulties and challenges using the MP intervention). In addition, participants completed a short questionnaire to evaluate the training sessions to identify issues using MP.

Sample size. The effect size of the primary outcome for effectiveness (APSPT 29) was used to cautiously estimate the sample size for a larger follow-up study as proposed by Dupont and Plummer[29] with the G*Power 3 application.[30]

For the power calculation of task procedure 1, a 2-tailed test with the error probability α of 0.05 and the power of 0.95 was set. These parameters predicted that a total sample size of 140 participants (ie, 70 in each group) would be required in a future study.

For the power calculation of task procedure 2, a 2-tailed test with the error probability α of 0.05 and the power of 0.95 was set.[30] These parameters predicted that a total sample size of 150 participants (ie, 75 in each group) would be required in a future study.

Data analysis

All statistical analyses were performed with the statistical programme R (version 3.4.3).[31] All data were first entered on a hard copy paper form that was subsequently entered electronically by a research assistant. The data set was checked for accuracy with a random check. Personal details were stored in a separate filing cabinet (ie, the master list with names was stored separately). Names in the database were replaced with an identification code.

Baseline characteristics of the participants (sex, age, primary language, and previous evaluation of procedural skills) were described overall and per intervention group.

Separate analyses were performed for the groups training the transfer procedure (task procedure 1) and the groups training the vestibular rehabilitation procedure (task procedure 2). Data were checked regarding normality to determine the appropriate analysis.[32] The evaluations indicated that the data were nonnormally distributed. Therefore, nonparametric tests, the Wilcoxon rank sum test,[33] were used to compare the performance between groups, and an effect size was presented for all analyses.[34] Effect sizes were interpreted as presented by Cohen:[35] an effect size (*r*) of 0.1 represented a small effect, a moderate effect was associated with 0.3, and a large effect was related to an effect size of 0.5 or more.

A distribution-based estimation of the minimally important difference was followed.[36] That is, we expressed the minimally important difference as 0.5 standard deviation.[37] For this calculation, we used the distribution of the scores at the post-acquisition endpoint. This resulted in 6.7 points for the APSPT score and 1.2 points for the PSC as meaningful changes.

The secondary outcomes and the different endpoints were analysed the same way. However, due to the large number of outcome measures, a correction for multiple tests was applied using the Benjamini and Hochberg[38] procedure. All analyses were performed as intention-to-treat analysis.[39] Missing data were obtained by performing multiple imputations using the Multivariate Imputation by Chained Equations package in R.[40] The number of imputations was set to 6.

The feasibility questionnaires were analysed by the first author using manual coding and thematic coding. Theme categories of problems and challenges were identified in the survey data and not prespecified. Consistency of coding was ensured by reanalysing the data in case of emergence of a new category.

Results

The participant flow for both study arms (task procedure 1: transfer procedure and task procedure 2: vestibular rehabilitation procedure) is presented below. The recruitment period started on September 1, 2017, and the trial was completed on January 31, 2018.

Participants – task procedure 1

A convenience sample of 18 participants was recruited for the study of task procedure 1. Participants were randomised as follows: group 1A 'MP' (n = 10); group 1B 'nMP' (n = 8).

All participants undertook the procedural skills training intervention and completed the post-acquisition test (T1) immediately after the intervention. The retention test (T2) was performed by 16 participants 2 weeks after the intervention. One participant was lost in group 1A and 1B, respectively (Figure 1). Both participants reported to be unavailable because of time constraints

Included participants are characterised in Table 1. Overall, considerably more women and French-speaking participants were included. However, this did not differ between groups. Previous academic performance was similar between groups. Marks in the Swiss educational system can range between 0 and 6 points. Higher scores indicate a better performance.

Participants – task procedure 2

A convenience sample of 19 participants was recruited for the task procedure 2 (vestibular rehabilitation procedure). Participants were randomly allocated as follows: group 2A 'mental practice (MP)' (n = 10) and group 2B 'no mental practice (nMP)' (n = 9). All included participants followed the procedural skills training intervention and completed the post-acquisition test immediately after the intervention. No participant was lost at the retention test (Figure 1).

Subject characteristics for participants are summarised in Table 2. Groups were similar regarding age and previous academic performance.

Analysis of effectiveness – task procedure 1

The findings of all outcome measures are presented in Table 3.

APSPT 29. At post-acquisition testing, the MP group had a higher median score of 82 (interquartile range [IQR]: 71-86) versus 74 (IQR: 70-78) for the nMP group. The difference was not statistically significant with a moderate effect size (*r*: −0.3) in favour of MP. At the retention test, both groups had reduced scores with the MP scored a median of 63 (IQR: 58-64) compared with 52 (IQR: 47-68) points in the nMP group. This difference represented a small effect (*r*: −0.17) but was not statistically significant (Figure 2).

Procedure-specific checklist. The MP group scored on average 1.5 points higher on a PSC (ranging from 0 to 24 points) at the post-acquisition test. A medium to large effect size was analysed for this difference (*r*: −0.42). This finding was not statistically significant.

The difference of medians: 1.5 points in favour of the MP group on the PSC remained at the retention test. This difference was not statistically significant with an effect size of *r*: −0.01.

Response time. The average response time was longer in the MP group compared with the nMP group (M: 263.6 [SD: 70.3] seconds versus 253 [SD: 37.19] seconds) at post-acquisition testing. This difference represented a moderate effect (*r*: −0.19) but was not statistically significant. A median response time of 295 seconds in the MP group compared with 282.5 seconds in the nMP was analysed. The effect size remained moderate (*r*: −0.19) and findings were statistically not significant.

Self-reported confidence. The nMP had a higher median score of 10 (IQR: 8-10) versus 9 (IQR: 8-9) points for the MP group at post-acquisition testing. This represented a small effect size (*r*: −0.18) but was not statistically significant.

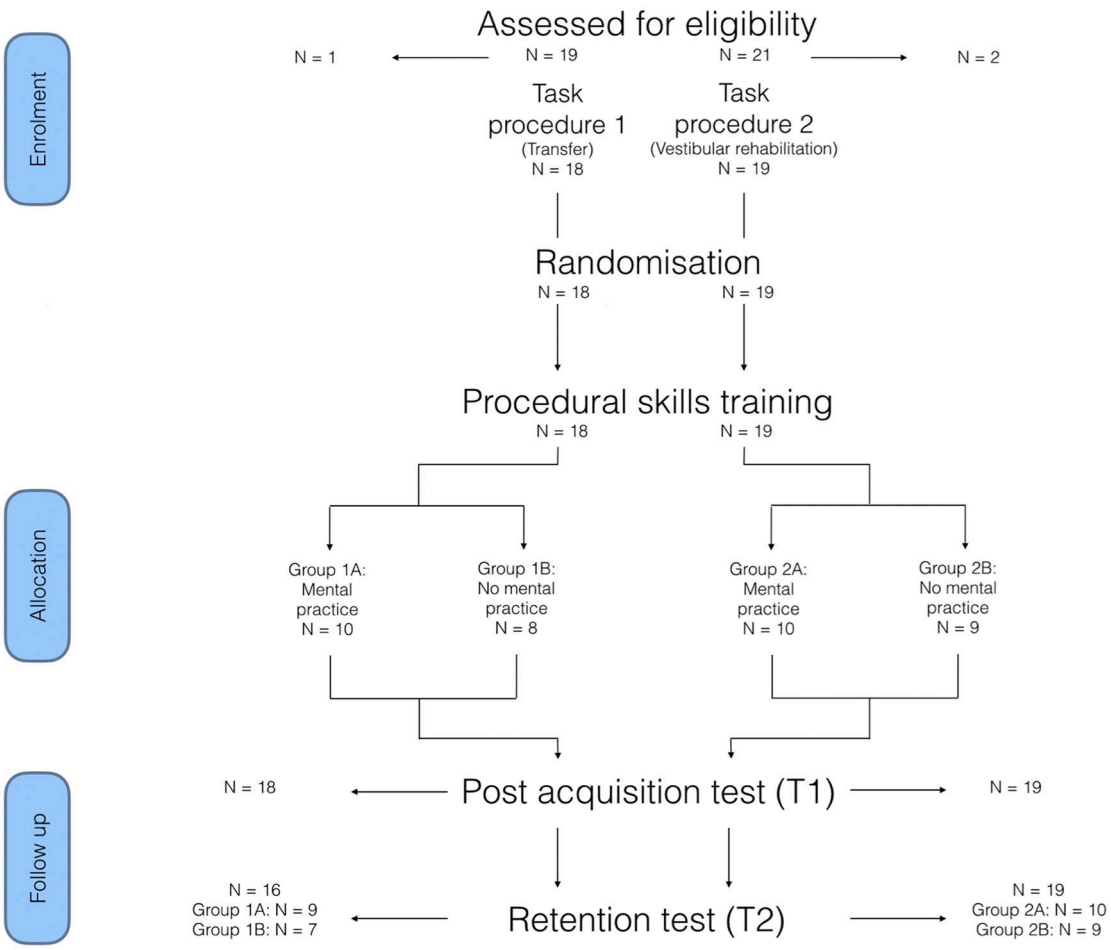


Figure 1. Study flow – task procedures 1 and 2 of the LEArN trial.
LEArN indicates Learning of Procedures in Physiotherapy Education.

Table 1. Demographic and educational data of included participants – task procedure 1.

VARIABLE	GROUP 1A (N=10)	GROUP 1B (N=8)	OVERALL
Sex			
Female	8	5	13
Male	2	3	5
Primary language			
Swiss French	8	5	13
Swiss German	2	3	5
Age (years)			
Mean (SD)	23.7 (1.85)	23.5 (1)	
Median (IQR)	23 (22-24)	23 (23-24)	
Previous examinations (range 0-6)			
Mean (SD)	5.1 (0.34)	5.1 (0.3)	
Median (IQR)	5 (5-5)	5 (4-5)	

Abbreviations: IQR, interquartile range.

Table 2. Demographic data of included participants – task procedure 2.

VARIABLE	GROUP 2A (N=10)	GROUP 2B (N=9)	OVERALL
Sex			
Female	8	8	16
Male	2	1	3
Primary language			
Swiss French	7	6	13
Swiss German	3	3	6
Age (years)			
Mean (SD)	26 (2.65)	25 (1.25)	
Median (IQR)	25 (24-27)	25 (24-25)	
Previous examination (range 0-6)			
Mean (SD)	4.93 (0.29)	4.96 (0.26)	
Median (IQR)	5 (5-5)	5 (5-5)	

Abbreviations: IQR, interquartile range.

Table 3. Effectiveness MP versus nMP – task procedure 1.

OUTCOME MEASURE		POST-ACQUISITION TEST (T1)		SIGNIFICANCE AND EFFECT SIZE	RETENTION TEST (T2)		SIGNIFICANCE AND EFFECT SIZE
		GROUP 1A (N = 10) MP	GROUP 1B (N = 8) NMP		GROUP 1A (N = 10) MP	GROUP 1B (N = 8) NMP	
APSPT 29 (0-116 points)	Median (IQR)	82 (71-86)	74 (70-78)	*W*: 55, *P*: .2; *r*: −0.3	63 (58-64)	52 (47-68)	*W*: 49, *P*: .45; *r*: −0.17
PSC (0-24 points)	Median (IQR)	21 (IQR: 20-23)	20 (19-20)	*W*: 61, *P*: .06 (.3*); *r*: −0.42	21 (IQ*R*: 19-22)	20 (19-23)	*W*: 39, *P*: .96 (.96*); *r*: −0.01
Response time (seconds)	Mean (SD)	263.6 (SD: 70.3)	253 (SD: 37.19)	*W*: 50, *P*: .41 (.72*); *r*: −0.19	300.2 (SD: 72.76)	284 (SD: 39.1)	*W*: 50, *P*: .4 (.7*); *r*: −0.19
Self-reported confidence (0-12 points)	Median (IQR)	9 (8-9)	10 (8-10)	*W*: 31, *P*: .43 (.72*); *r*: −0.18	7 (6-9)	9 (8-10)	*W*: 25.5, *P*: .21 (.7*); *r*: −0.29

Abbreviations: APSPT 29, Assessment for Procedural Skills in Physiotherapy Education with 29 items; IQR, interquartile range; MP, mental practice; nMP, no mental practice; PSC, procedure-specific checklist; *r*, effect size; *W*, Wilcoxon rank sum statistic.
**P* value corrected for multiple testing.

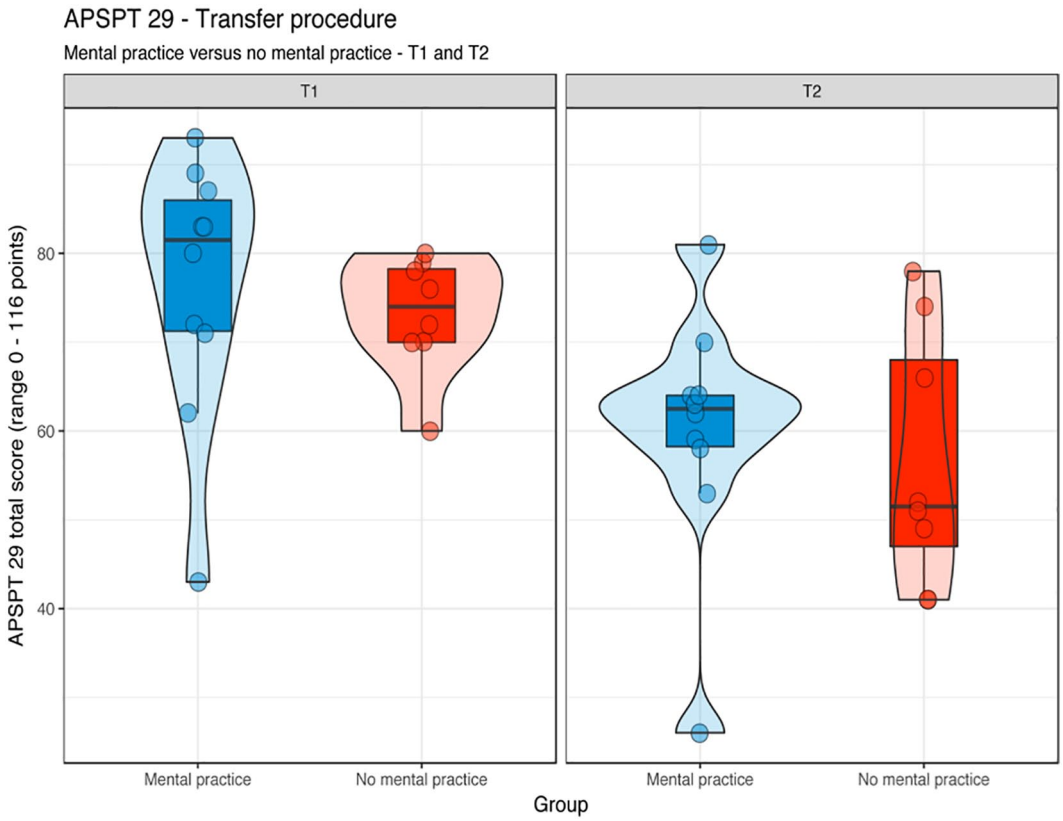


Figure 2. APSPT 29 – comparison of MP (1A) versus nMP (1B) task procedure 1. Individual participant data are presented as dots. Density curves are plotted in blue (MP) or red (nMP).
APSPT indicates Assessment for Procedural Skills in Physiotherapy Education with 29 items; MP, mental practice; nMP, no mental practice; T1, post-acquisition test; T2, retention test.

A median of 7 (IQR: 6-9) points in the MP group versus 9 (IQR: 8-10) in the nMP group were recorded at retention testing. The difference was not statistically significant and represented a medium effect size in favour of the nMP group (*r*: -0.29).

Analysis of the feasibility – task procedure 1

Recruitment rate. Nineteen participants were invited to participate in this study, with a total of 18 being recruited. This resulted in a recruitment rate of 95%.

Table 4. Effectiveness MP versus nMP – task procedure 2.

OUTCOME MEASURE		POST-ACQUISITION TEST (T1)		SIGNIFICANCE AND EFFECT SIZE	RETENTION TEST (T2)		SIGNIFICANCE AND EFFECT SIZE
		GROUP 2A (N = 10) MP	GROUP 2B (N = 9) NMP		GROUP 2A (N = 10) MP	GROUP 2B (N = 9) NMP	
APSPT 29 (0-116 points)	Median (IQR)	74 (61-83)	66 (54-69)	*W*: 61, *P*: .21; *r*: −0.29	63 (45-86)	57 (55-84)	*W*: 45.5, *P* > .99; *r*: 0
PSC (0-26 points)	Median (IQR)	21 (18-22)	19 (IQR: 18-21)	*W*: 57, *P*: .34 (.43*); *r*: −0.22	20 (15-23)	20 (13-22)	*W*: 46.5, *P*: .93 (.99*); *r*: −0.01
Response time (seconds)	Mean (SD)	199.8 (SD: 75.58)	150.89 (SD: 37.85)	*W*: 64, *P*: .13 (.21*); *r*: −0.34	263.9 (SD: 65.48)	216.22 (SD: 67.96)	*W*: 67, *P*: .08 (.23*); *r*: −0.39
Self-reported confidence (0-12 points)	Median (IQR)	9 (6-10)	5 (5-8)	*W*: 69, *P*: .049 (.21*); *r*: −0.45	8 (7.25-9)	8 (7-9)	*W*: 48.5, *P*: .8 (.99*); *r*: −0.06

Abbreviations: APSPT 29, Assessment for Procedural Skills in Physiotherapy Education with 29 items; IQR, interquartile range; MP, mental practice; nMP, no mental practice; PSC, procedure-specific checklist; *r*, effect size; *W*, Wilcoxon rank sum statistic.
**P* value corrected for multiple testing.

Failure rate. At the post-acquisition endpoint, 1 MP participant was classified as 'failed to perform' the procedure. This equates to a 10% failure rate in the MP group. This was considerably below the predefined threshold of 40%. At the retention test, failures were observed in both groups. This resulted in failure rates of 11.1% (1 of 9) in the MP and 14.3% (1 of 7) in the nMP group

Feasibility of the procedural skills training session. Within the protocol, it was assumed that the MP intervention might be difficult to use for the procedural skills training as it is not a standard intervention. Participants were given explanation regarding how to perform the MP. During the practice, no further questions were asked and on checking, all participants (ie, the students) indicated that they could perform the MP instructions. Below, the key data from the analysis of the surveys are reported. The process of thematic coding resulted in the emergence of 8 theme categories of problems and challenges: *communication, ergonomic aspects, hand placement, memory, peer learning, preparation, security, simulation, and specific procedure parts*. Most survey responses were tagged with 1 theme, but this ranged up to 3 themes per survey response. All theme categories of identified challenges and corresponding frequencies are reported in Appendix 2. The most frequently reported challenges in the MP group were related to (1) the hand placement during the procedure and (2) memorisation processes. Specifically, it was mentioned that the relatively large number of cues was challenging to memorise. Most challenges in the nMP group were related to (1) ergonomic aspects of the procedure and (2) the hand placement and the security of the patient.

Analysis of effectiveness – task procedure 2

The findings of all outcome measures are presented in Table 4.

APSPT 29. At post-acquisition testing (T1), the MP group had a higher median score of 74 (IQR: 61-83) versus 66 (IQR: 54-69) for the nMP group on the APSPT 29 (possible range between 0 and 116 points). This difference was not statistically significant. The effect size of *r*: −0.29 indicated a moderate effect size in favour of the MP group. At the retention test (T2), MP had a slightly higher median score than nMP (63 [IQR: 45-86] versus 57 [IQR: 55-84] points). The difference was not statistically significant.

Procedure-specific checklist. The median performance measured with a PSC (possible range between 0 and 26 points) at the post-acquisition test was higher in the MP group (median: 21 [IQR: 18-22] points) compared with the nMP group (median: 19 [IQR: 18-21] points). A small to moderate effect size of *r*: −0.22 was analysed in favour of the MP group. Findings were statistically not significant. The point estimates at the retention test were similar between groups, and the between-group difference was not statistically significant.

Response time. The mean response time (ie, the duration of the procedure) was longer in the MP group (199.8 [SD: 75.58] seconds) compared with the nMP group (150.89 [SD: 37.85] seconds) at the post-acquisition test. The effect size for this outcome measure was moderate (*r*: −0.34) but was not statistically significant. This finding did not change at the retention test. The mean duration was 263.9 (SD: 65.48) seconds in the MP group versus 216.22 (SD: 67.96) seconds in the nMP group. The effect size for this between-group difference was moderate to large (*r*: −0.4) but the finding was not statistically significant.

Self-reported confidence. The self-reported confidence in the MP group was higher (median: 9 [IQR: 6-10] points) compared with 5 (IQR: 5-8) points in the nMP group at the post-acquisition test

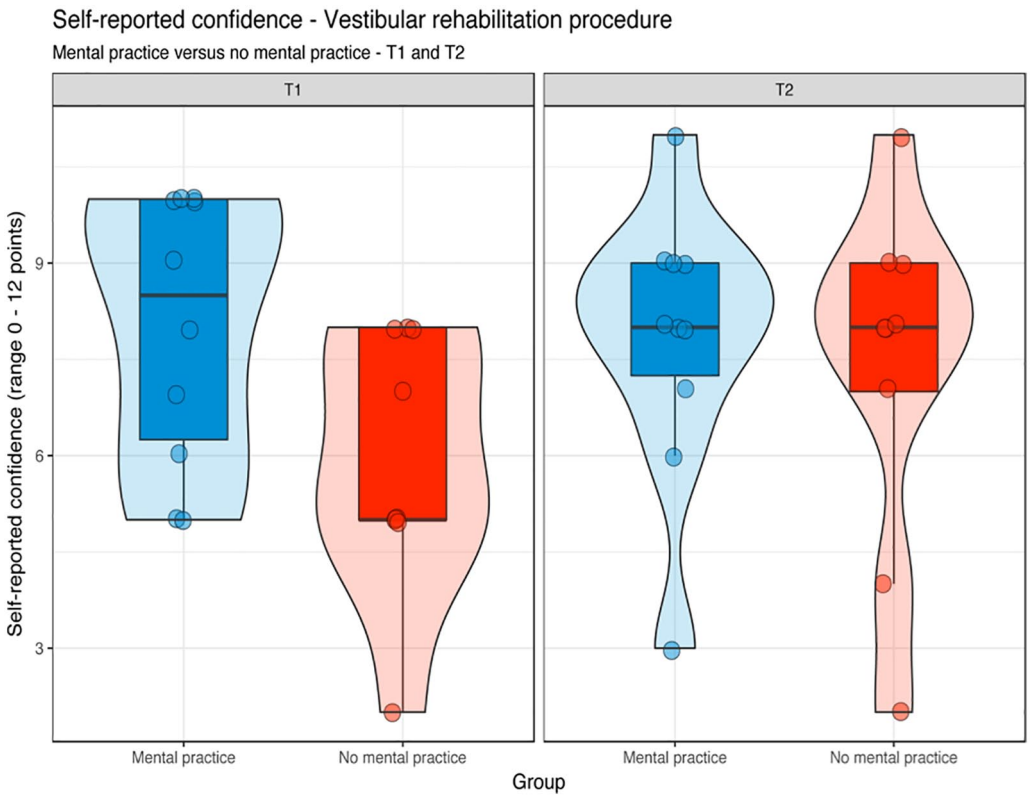


Figure 3. Self-reported confidence – comparison of MP (2A) versus nMP (2B) task procedure 2. Individual participant data are presented as dots. Density curves are plotted in blue (MP) or red (nMP).
MP indicates mental practice; nMP, no mental practice; T1, post-acquisition test; T2, retention test.

(Figure 3). The effect size of this between-group difference was moderate to large (*r*: -0.45) and statistically significant (*P*: 0.049). However, after correction for multiple testing, the *P* value increased to .21, and the findings were above the significance level. At the retention test, both groups had a similar self-reported confidence.

Analysis of the feasibility – task procedure 2

Recruitment rate. Twenty-one participants were approached for this study, with a total of 19 being recruited. This resulted in a recruitment rate of 91%.

Failure rate. Testing of participants, post-acquisition, showed only 1 MP participant classified as 'failed to perform'. This equates to a 10% failure rate in the MP group. This was considerably below the predefined threshold of 40%. In the nMP group, all participants adequately performed the procedure post-acquisition. At the retention test, the number of participants 'failing to perform' increased in both groups. This resulted in failure rates of 30% (3 of 10) in the MP group and 37.5% (3 of 8) in the nMP group.

Feasibility of the procedural skills training session. The use of the MP provided some challenges as it was unfamiliar for participants. If the participants had questions regarding the MP, these were answered. Despite the novel situation, all participants performed the MP.

Below, the key data from the analysis of the surveys are reported. The process of thematic coding resulted in the emergence of 10 theme categories of problems and challenges: *communication, concentration, hand placement, memory, MP training, preparation, reasoning, security, video assessment, and visualisation vestibular system*. Most survey responses were tagged with 1 theme, but this ranged up to 3 themes per survey response. All theme categories of identified challenges and corresponding frequencies are reported in Appendix 3. Most reported challenges in the MP group were related to (1) MP was a novel kind of practice not used previously, (2) the large number of procedural steps caused recall problems, and (3) three participants mentioned MP was difficult because of a noisy environment. A further challenge was the visualisation of the vestibular system. Most challenges in the nMP group were related to the placement of the hands during the execution of the procedures.

Discussion

The main finding of this comparison was that a moderate effect size in favour of MP was found on the APSPT 29 for both task procedures at post-acquisition testing. The observed difference in medians was in favour of a procedural skills training with MP compared with a training using only physical practice (8 and 7.5 points difference for task procedures 1 and 2, respectively). Using 6.7 points (based on a distribution-based estimation) as an indicator of the minimally important difference indicates that MP can potentially be used to increase

acquisition of procedural skills. However, because the findings were not statically significant, it is not possible to state that MP was more effective for skill acquisition of procedures in physiotherapy education than nMP.

The same possible trend was found on a PSC. Moderate effect sizes identified were in favour of the MP group. However, the potential effect of MP on a PSC was smaller. It was estimated that 1.2 points would indicate a minimally important difference and the difference in medians was close but smaller for both task procedures.

These findings are supported by evidence from other studies in health professions education. For example, a meta-analysis published by Sattelmayer et al[15] found a moderate effect size of 0.43 SMDs in favour of an MP intervention at post-acquisition testing. Effect sizes for individual studies ranged from −0.09[9] to 1.80 SMD.[41] Furthermore, of the five studies included into the meta-analysis, only one study showed statistically significant results and a high effect size.[41] The effect size of the current study was considerably lower and more in line with the findings of the remaining four studies.[9,24,42,43]

The observed point estimates and the effect sizes of the APSPT 29 at the retention test remained in favour of the MP groups. However, effect sizes and the difference in medians were smaller but the between-group difference was still in the range of a minimally important difference. Only 2 studies were found using a transfer test to make inferences on learning in health professions education.[9,24] Both studies showed a small effect size in favour of MP at transfer testing. This is similar to the findings of this study.

The findings from the post-acquisition endpoint and retention test indicate that MP could provide a meaningful difference on the acquisition of procedural skills in physiotherapy education. Given that the sample was underpowered, it is essential to confirm these preliminary results with an adequately powered follow-up study. One important finding was related to the participant's self-reported confidence. It was expected that a higher self-confidence would be associated with a better performance. At least for athletic performance, there seems to exist such an association.[44] For task procedure 1, there was no such association. However, the self-reported confidence of the MP group training task procedure 2 was higher with a moderate to large effect size. This suggests that MP could be used to increase self-confidence in one's own abilities.

Regarding the outcome movement duration, it was observed that the median response time was longer in the MP groups. A priori it was expected that with an increased performance, a reduction in response time would be associated. For example, Starkes et al[23] reported that the time needed to perform a procedure in surgery was a good estimate of the individual's procedural abilities. However, in this study, the MP groups (ie, the group with the higher median performance ratings) needed considerably more time than the nMP groups at both endpoints. Two reasons might have caused the longer response time: (1) nonmotor aspects (such as an increased amount of information provided) and (2) a longer patient assessment. Providing instructions and post-procedure information was rated on both performance indices (APSPT 29 and PSC). A detailed patient instruction, which requires time, did therefore increase the performance rating.

Based on these findings, it can be hypothesised that increased procedural competency is not inevitably associated with a decreased response time in these complex manual procedures. Furthermore, the outcome 'response time' should be evaluated cautiously in physiotherapy education and the assumption that a shorter response time indicates proficiency might not be valid in this setting.

Discussion of feasibility

The main finding of the feasibility analysis for the MP against nMP comparison was that the feasibility of the study was high based on quantitative feasibility measures.

The observed recruitment rate was very high, and the failure rate was very low at the post-acquisition endpoint; although the failure rate increased at the retention test, both groups remained below the predefined 40% threshold. The increased failure rate at the retention test indicates that genuine learning did not occur in all participants, and for parts of the sample, more practice would be required.

A key feasibility issue noted was that some participants felt challenged with the MP intervention. Mental practice is not a standard intervention, and it is difficult to use a new training principle for the first time. Some participants stated that it was difficult to perform an MP of the procedures because of the surrounding noise. In a future study, this should be addressed. For example, all participants should be reminded to be calm, or be offered private practice rooms.

Another aspect might have decreased the feasibility of the MP training. Within this study, a ratio of 1 practice trial to 1 mental rehearsal trial was instructed. This was based on the best practice recommendations for mental imaging published by Schuster et al.[11] However, other authors recommend to gradually increase the amount of mental repetitions between physical practice trials. Initially, the authors recommend to start with a ratio of 1 physical practice trial to 5 mental repetitions, which may be increased to a ratio of 10 mental to 1 physical practice repetitions with increasing proficiency.[45] This might have increased the ability to independently use MP for skill acquisition.

Limitations

This study was designed as a pilot study intended to investigate a new method in physiotherapy education. With the recruited sample size, it was possible to estimate effect sizes and investigate whether the between group difference was large enough to be meaningful. However, the sample size of this study was underpowered to confirm the findings with statistical significance testing.

In contrast to other studies, eg, Arora et al[18] and Sanders et al,[24] no relaxation exercises were performed prior to the exercises, which might have further increased the feasibility (ie, the participants would be more prepared to use the technique) and effectiveness of the MP intervention. This approach was not followed because of pragmatic reasons. Malouin et al[46] report that the use of relaxation exercises may be associated with potential benefits such as increased concentration and attention, a more vivid mental imaging and an increased motor performance. Especially, the study of Arora et al[41] showed that MP was more effective than physical practice alone. This might be caused by their use of relaxation exercises prior to the MP. Therefore, adding relaxation exercises to the MP intervention might increase the effectiveness in future studies, and the lack of relaxation exercises might be regarded as a limitation of the LEArN trial.

An additional limitation of this study was that the used outcome measures were not fully validated in advance. Only limited evidence is available regarding the measurement properties of the APSPT 29. Evidence for the following measurement properties was missing: reliability (test-retest, intrarater, and interrater), measurement error, criterion validity, construct validity, and responsiveness. To partly address these limitations, the interrater reliability of APSPT 29 was analysed. A second independent rater was asked to score the video recordings of the retention test. An intraclass correlation coefficient (ICC; 2,1) of 0.79 for the total score of the APSPT 29 was analysed indicating adequate interrater reliability. The same limitations apply to the PSC. A secondary analysis showed an ICC (2,1) of 0.92 for the total score of the PSC indicating adequate interrater reliability. However, these analyses were performed after the primary analyses of the LEArN trial, which is a considerable limitation.

In addition, a modified questionnaire was used to establish self-reported confidence, which was not validated in a previous study and the used feasibility surveys were not validated in a pilot study.

Finally, ethical implications are associated with this study design. The primary investigator and the participants were all affiliated with the UAS Valais and a power difference exists, which might have caused participants to feel obligated to participate. This cannot be excluded but several precautions were taken. For example, the study information was provided by an independent person, and participants were informed that their performance would be rated by an external person not involved in teaching at the UAS Valais. The proposed measures to avoid such a feeling were presented to the ethical committee and ethical clearing was granted. In addition, from a qualitative research paradigm, an insider's view has also advantages such as immediate legitimacy in the field and a nuanced perspective for observation, interpretation, and representation.[47]

Conclusions

Based on this pilot study, the effectiveness of MP on skill acquisition in physiotherapy education could be estimated. Moderate effect sizes in favour of MP were identified at post-acquisition testing, and at retention testing, the between-group differences were considerably smaller. Furthermore, the between-group difference at post-acquisition test represented a meaningful difference. However, the results were not statistically significant for the transfer procedure and the vestibular rehabilitation procedure. The use of MP in a physiotherapeutic educational setting is feasible and should be further explored in follow-up studies.

Recommendations for research are (1) to evaluate the effectiveness of MP versus nMP in future studies with a larger sample size and (2) to analyse the effectiveness of MP on other procedures in physiotherapy education.

Authors' Note

The LEArN trial was part of the doctoral thesis of the first author (K.M.S.) at Queen Margaret University in Edinburgh: Sattelmayer KM. *Acquisition of complex therapeutic procedures in pre-registration physiotherapy education using motor learning principles*, Queen Margaret University, Edinburgh; 2019.
The authors have full rights to publish the manuscript.

Author Contributions

All authors have made substantial contributions to the conception of the study and have approved the submitted version of the study. In addition, KMS has been involved in the conceptual development, design, analysis and the interpretation of data. KCJ has been involved in the conceptual development, design, interpretation of data and has substantively revised the manuscript. RH has been involved in the analysis of data and the interpretation of data. GB has been involved in the conceptual development, design, analysis and interpretation of data and has substantively revised the manuscript

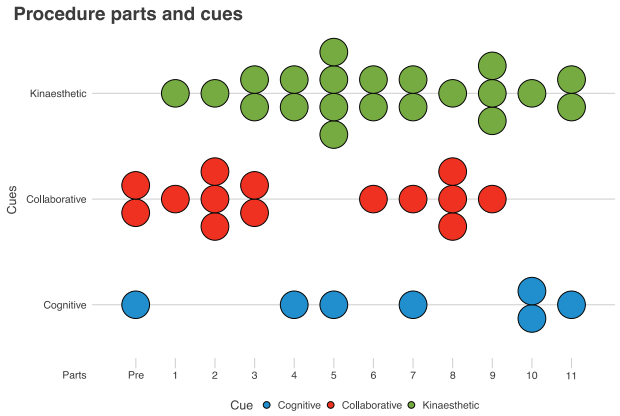
ORCID iD

Karl Martin Sattelmayer https://orcid.org/0000-0001-8001-4776

REFERENCES

1. World Confederation for Physical Therapy. Policy statement: education. https://www.wcpt.org/policy/ps-education. Updated 2017.
2. Gorrell LM, Engel RM, Brown B, Lystad RP. The reporting of adverse events following spinal manipulation in randomized clinical trials – a systematic review. *Spine J*. 2016;16:1143-1151.
3. Maran NJ, Glavin RJ. Low-to high-fidelity simulation–a continuum of medical education? *Med Educ*. 2003;37:22-28.
4. Reed S, Shell R, Kassis K, et al. Applying adult learning practices in medical education. *Curr Probl Pediatr Adolesc Health Care*. 2014;44:170-181.
5. Faustinella F, Jacobs RJ. The decline of clinical skills: a challenge for medical schools. *Int J Med Educ*. 2018;9:195-197.
6. Kelly M, Lyng C, McGrath M, Cannon G. A multi-method study to determine the effectiveness of, and student attitudes to, online instructional videos for teaching clinical nursing skills. *Nurse Educ Today*. 2009;29:292-300.

7. Rourke S. How does virtual reality simulation compare to simulated practice in the acquisition of clinical psychomotor skills for pre-registration student nurses? A systematic review. *Int J Nurs Stud*. 2019;102:103466.
8. Jang HW, Kim K-J. Use of online clinical videos for clinical skills training for medical students: benefits and challenges. *BMC Med Educ*. 2014;14:56.
9. Sanders CW, Sadoski M, van Walsum K, Bramson R, Wiprud R, Fossum TW. Learning basic surgical skills with mental imagery: using the simulation centre in the mind. *Med Educ*. 2008;42:607-612.
10. Schmidt RA, Lee T. *Motor Control and Learning 5e*. Champaign, IL: Human Kinetics; 2011.
11. Schuster C, Hilfiker R, Amft O, et al. Best practice for motor imagery: a systematic literature review on motor imagery training elements in five different disciplines. *BMC Med*. 2011;9:75.
12. Landers DM. The effects of mental practice on motor skill learning and performance: a meta-analysis. *J Sport Psychol*. 1983;5:25-57.
13. Feltz DL, Landers DM, Becker BJ. *A Revised Meta-Analysis of the Mental Practice Literature on Motor Skill Learning*. Washington, DC: National Academy Press; 1988.
14. Wohldmann EL, Healy AF, Bourne LE Jr. A mental practice superiority effect: less retroactive interference and more transfer than physical practice. *J Exp Psychol Learn Mem Cogn*. 2008;34:823-833.
15. Sattelmayer M, Elsig S, Hilfiker R, Baer G. A systematic review and meta-analysis of selected motor learning principles in physiotherapy and medical education. *BMC Med Educ*. 2016;16:15.
16. Dobkin BH. Progressive staging of pilot studies to improve phase III trials for motor interventions. *Neurorehabil Neural Repair*. 2009;23:197-206.
17. Snow G. blockrand: randomization for block random clinical trials. R package version 1.3. https://cran.r-project.org/web/packages/blockrand/index.html. Updated 2013.
18. Arora S, Aggarwal R, Sevdalis N, et al. Development and validation of mental practice as a training strategy for laparoscopic surgery. *Surg Endosc*. 2010;24:179-187.
19. Ericsson KA. Protocol analysis and expert thought: concurrent verbalizations of thinking during experts' performance on representative tasks. In: Ericsson KA, Charness N, Feltovich PJ, Hoffman RR, eds. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge, UK: Cambridge University Press; 2006:223-242.
20. Cumming J, Olphin T, Law M. Self-reported psychological states and physiological responses to different types of motivational general imagery. *J Sport Exerc Psychol*. 2007;29:629-644.
21. Sattelmayer KM. Development and validation of a mental practice script for a transfer procedure for people with hemiparesis after stroke. In: Sattelmayer KM ed. *Acquisition of Complex Therapeutic Procedures in Pre-Registration Physiotherapy Education Using Motor Learning Principles*. Edinburgh, UK: Queen Margaret University; 2019:79-94.
22. Sattelmayer M, Hilfiker R, Baer G. Développement d'une échelle d'évaluation des compétences procédurales dans la formation en physiothérapie: une étude de validation de la théorie de la réponse par item. *Kinésithérapie, la Revue*. 2016;16:38.
23. Starkes JL, Payk I, Hodges NJ. Developing a standardized test for the assessment of suturing skill in novice microsurgeons. *Microsurgery*. 1998;18:19-22.
24. Sanders CW, Sadoski M, Bramson R, Wiprud R, Van Walsum K. Comparing the effects of physical practice and mental imagery rehearsal on learning basic surgical skills by medical students. *Am J Obstet Gynecol*. 2004;191:1811-1814.
25. Edwards J, Sanders C, Sadoski M, van Walsum K, Bramson R, Wiprud R. The effect of imagery rehearsal on medical student learning of basic surgical skills, in a collaborative environment. Paper presented at: Annual Meeting of the American Educational Research Association; 21-25 April 2003, Chicago, IL.
26. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol*. 2010;10:1.
27. Van Teijlingen E, Hundley V. The importance of pilot studies. *Nurs Stand*. 2002;16:33-36.
28. Ten Cate O, Regehr G. The power of subjectivity in the assessment of medical trainees. *Acad Med*. 2019;94:333-337.
29. Dupont WD, Plummer WD Jr. Power and sample size calculations: a review and computer program. *Control Clin Trials*. 1990;11:116-128.
30. Faul F, Erdfelder E, Lang A-G, Buchner A. G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39:175-191.
31. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2014.
32. Field A. *Discovering Statistics Using IBM SPSS Statistics*. London, England: SAGE; 2013.
33. Wilcoxon F. Individual comparisons by ranking methods. *Biom Bull*. 1945;1:80-83.
34. Rosenthal R, Rosnow RL, Rubin DB. *Contrasts and Effect Sizes in Behavioral Research: A Correlational Approach*. Cambridge, UK: Cambridge University Press; 2000.
35. Cohen J. A power primer. *Psychol Bull*. 1992;112:155-159.
36. Rai SK, Yazdany J, Fortin PR, Aviña-Zubieta JA. Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. *Arthritis Res Ther*. 2015;17:143.
37. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-592.
38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995:289-300.
39. Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol*. 1992;21:837-841.
40. Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1-67.
41. Arora S, Aggarwal R, Sirimanna P, et al. Mental practice enhances surgical technical skills: a randomized controlled study. *Ann Surg*. 2011;253:265-270.
42. Geoffrion R, Gebhart J, Dooley Y, et al. The mind's scalpel in surgical education: a randomised controlled trial of mental imagery. *BJOG*. 2012;119:1040-1048.
43. Komesu Y, Urwitz-Lane R, Ozel B, et al. Does mental imagery prior to cystoscopy make a difference? A randomized controlled trial. *Am J Obstet Gynecol*. 2009;201:218.e1-218.e9.
44. Feltz DL. Self-confidence and sports performance. *Exerc Sport Sci Rev*. 1988;16:423-458.
45. Malouin F, Richards CL. Mental practice for relearning locomotor skills. *Phys Ther*. 2010;90:240-251.
46. Malouin F, Jackson PL, Richards CL. Towards the integration of mental practice in rehabilitation programs. *Front Hum Neurosci*. 2013;7:576.
47. Chavez C. Conceptualizing from the inside: advantages, complications, and demands on insider positionality. *Qual Rep*. 2008;13:474-494.



Appendix 1 Overview MP script for task procedure 1.
MP indicates mental practice.
The transfer procedure (task procedure 1) was appraised as having 1 preparation phase (PRE) and 11 procedure parts. For each procedure part, cognitive (blue), collaborative (red), and kinaesthetic cues (green) are presented. A specific example is given for procedure part 5. The task for this procedure part was 'the patient is asked to sit on one side'. The cognitive cue was 'Make the decision to go over the weaker or the stronger side'. A kinaesthetic cue for this procedure part was 'The patient slowly slides over your legs to the floor'.
The MP script for task procedure 2 was designed using the same structure.

Appendix 2 Feasibility analysis survey data task procedure 1.

THEME CATEGORIES OF CHALLENGES AND PROBLEMS	COUNT OF CATEGORIES IN GROUP 1A (MP)	COUNT OF CATEGORIES IN GROUP 1B (NMP)
Communication	1	0
Ergonomic aspects	0	2
Hand placement	3	2
Memory	3	0
Peer learning	0	1
Preparation	1	0
Security	1	2
Simulation	1	0
Specific procedure parts	2	0

Abbreviations: MP, mental practice; nMP, no mental practice.

Appendix 3 Feasibility analysis survey data task procedure 2.

THEME CATEGORIES OF CHALLENGES AND PROBLEMS	GROUP 2A (MP)	GROUP 2B (NMP)
Communication	0	1
Concentration	3	0
Hand placement	0	3
Memory	2	2
MP training	4	0
Preparation	0	1
Reasoning	1	2
Security	0	2
Video assessment	1	0
Visualisation vestibular system	2	2

Abbreviation: MP, mental practice; nMP, no mental practice.