Tailoring to educational needs: preparatory studies into doctor-patient communication training and the development of trainers’ expertise in general practice specialty training

van Es, J.M.

Citation for published version (APA):
van Es, J. M. (2012). Tailoring to educational needs: preparatory studies into doctor-patient communication training and the development of trainers’ expertise in general practice specialty training.
CHAPTER 6

SMART, SMARTER, SMARTEST: The influence of peer groups compared to practice visits on the quality of action plans

Judy van Es, Mechteld Visser, Margreet Wieringa-de Waard

Abstract

**Background**
It has been reported that appraisal by peers can be effective.

**Aim**
To investigate whether feedback from a peer group (PG) compared to feedback by a staff member during a practice visit (PV) is as effective in improving the quality of action plans.

**Methods**
73 GP trainers randomized into either a PG or a PV, were instructed to draw up action plans using the SMART-criteria, to realize the goals set in their Personal development Plans (PDP's). To improve action plans feedback was given in either PG or PV. Quality of baseline and follow-up action plans, operationalized as the SMARTness with which plans were formulated, was assessed using a study-specific instrument.

**Results**
Response rate for submitting both baseline and follow-up action plans was 89% in the PG versus 79% in the PV. It was feasible to determine scores on all SMART-criteria, except for the criterion “Acceptability”. Significant improvement was made on the remaining four criteria irrespective of the feedback setting.

**Conclusions**
PGs, cost less and seem equally effective in improving the SMARTness of action plans. Moreover, they also seem to stimulate GP trainers more to write a PDP. Therefore they may be favoured over PVs.
Introduction

Self-directed learning is a process in which individuals take the initiative to diagnose their learning needs, design learning experiences, locate resources and evaluate their progress. The concept has been applied by many educators the last three decades. To promote the self-directed learning processes and monitor their results, portfolios were introduced. These have proven to be effective in making learners more responsible for their own learning.

One item very often included in a portfolio is the personal development plan (PDP), also referred to as a personal learning plan. PDPs are usually the result of a structured formative assessment process that involves collecting and reviewing external feedback, as well as guided self-reflection activities. Self-reflection combined with external feedback enables learners to define areas in which further (educational) development is needed and to translate these needs into learning goals.

Commonly, in a PDP the approach to achieve developmental goals is formulated by way of action plans. A common tool used to help improving the quality of an action plan in profit- and non-profit organisations like the educational setting is the SMART-acronym. SMART stands for 5 criteria: Specific, Measurable, Acceptable, Realistic and Time-bound (sometimes slightly different terms are used for this acronym: attainable instead of acceptable, for instance). These criteria are meant to help define goals and formulate action plans clearly and enhance their effectiveness. As such, the SMARTness of an action plan can be seen as an aspect of an action plan’s quality.

In our context, General Practitioner (GP) speciality training, the clinical teachers, GP trainers, play a central role in training future GPs. GP trainers (experienced GPs) are required to follow modular GP trainers’ courses to keep their knowledge and teaching skills up to date. To monitor and manage the performance of individual trainers, practice visits are performed by the teaching staff (GPs and behavioural scientists) of the GP speciality training.

But unlike the standardized procedure that is generally followed for practice visits as part of the appraisal of GP practices, our educational practice visits to our GP trainers were until recently not standardized. O’Brien concluded that (educational) standardized practice visits that make use of specific tools to assess developmental needs have an effect on professional practice and Finlay...
Tailoring to educational needs

found that GPs consider them to enhance their learning and improve their practice.\textsuperscript{9,10} In line of these findings we recently implemented a format for our practice visits that standardizes both the interview during the visit and the subsequent reporting of this interview. We also asked our GP trainers to write a PDP before receiving a practice visit, herewith introducing a specific tool to enable us to gain a better insight into their developmental needs. We helped them to define goals by providing them with feedback consisting of their scores on a self-assessment questionnaire and their scores on the evaluations of their GP trainees of the last four years.\textsuperscript{11} GP trainers used the feedback for more than 75\% of the goals they defined.

In this study we will focus on GP trainers’ goal and action plan-writing abilities by investigating how SMART action plans for the goals in their PDPs are defined - the SMARTness of the action plans. Assuming that the SMARTer an action plan is defined, the greater the chances are of achieving the goal, this should ultimately lead to GP trainers managing their developmental needs more adequately.

Practice visits, however, take up a lot of staff time. Practices are located within a 100-kilometre radius of the GP Speciality Training institute where the teaching staff is based. As a consequence, in addition to preparation time and time for writing reports on the practice visits, much time is spent travelling.

In literature on GP practice appraisal it has been reported that appraisal by peers can be just as effective as appraisal by trained non-peers.\textsuperscript{12,13} Furthermore, feedback is considered to be more worthwhile when it has been given by peers.\textsuperscript{12,13} Obviously, peer groups take considerably less staff time: they address multiple GP trainers at once and reduce travel time. As such they could be a cost-effective way to manage GP trainers’ individual developmental needs than practice visits.

Considering the above we will investigate whether the effect of feedback on the SMARTness of action plans from peer groups equals that of feedback in a practice visit.

Herefore we need an instrument to assess the SMARTness of an action plan. Since, to the best of our knowledge, no such instrument exists, we will first develop one.
Methods

Participants
This study was conducted in 2008-2009 at the Speciality Training for General Practice of the Academic Medical Center (AMC), University of Amsterdam. All GP trainers of 1st year GP trainees in 2008 with at least two years’ experience as a GP trainer (n=85) were invited to participate in the study. GP speciality training is a three-year post-graduate training programme. In year one and three trainees are allocated to a GP trainer for the entire year. They also attend modular courses at the GP training institute one day a week. In the second year they work through rotations in clinical settings.

Study design and procedures
We performed a randomized trial. The SMARTness of the action plans was assessed before intervention (baseline) and after intervention (follow-up). GP trainers were randomised into either of the two intervention groups: peer-group meeting or practice visit. Subsequently, to help GP trainers choose relevant goals for their PDPs, we provided them with an overview of written competency-based feedback. This feedback consisted of self-assessment scores and the ratings scores and narrative comments from the evaluations of the trainees who had been allocated to their practice the last four years. Trainers were asked to write a PDP (baseline) using the feedback provided. We did not limit the number of goals that could be addressed by action plans in a PDP. At the same time, a date was set for either participation in a peer group, or a practice visit. During the peer-group meetings or practice visits, they received feedback on the SMARTness of their PDP’s. For each GP trainer, the first goal and accompanying action plan was used for this study. They were subsequently asked to review and adjust their action plans using the feedback they received. Both the initial (baseline) and the definitive (follow-up) action plans were sent to a research assistant to be anonymised before their SMARTness was scored by the first and the last author (JVE, MW). Peer groups comprised between three and five GP trainers and teaching staff of the GP training institute presided them. GP trainers could sign up for different dates and groups were compiled accordingly. Peer group meetings were held at the GP training institute and some (to reduce travel time for GP trainers) at the GP practice of one of the participating GP trainers. GP trainers who worked
Tailoring to educational needs

at the same practice were not allocated to the same peer group. Peer-group meetings were scheduled to last 3 hours.
Practice visits were also carried out by teaching staff and lasted 1.5 hours. For both peer groups and practice visits we made sure teaching staff having other professional or private relationships with a GP trainer were not assigned to that particular trainer. Two weeks prior to peer-group meetings or practice visits, presiding teaching staff received the feedback overviews and baseline PDPs for preparation purposes.

Instruction of participants
All participants were informed about this study during a meeting on a modular-course day. At this meeting, the first author explained the principles of writing a PDP. Participants were shown the layout of the overview in which their feedback would be collected. They were urged to use the feedback in drawing up their PDP. Participants were further instructed to structure their goals in action plans by using the subheadings aim, action, result and time span (AART) and they were encouraged to use the SMART-criteria to evaluate their action plans.
Participation was part of the regular modular course and therefore obligatory. Each participant gave written informed consent for using his or her PDP for this study. GP trainers who did not submit PDPs were reminded to do so several times by both email and telephone.
Data on the personal characteristics of participants were gathered from the administrative database of our institute. The ethics committee of the Dutch Society of Medical Education (NVMO) approved this study and exempted it from judgement by the Medical Ethical Committee since it involves no patients.

Instruction of teaching staff
All teaching staff at our GP training institute is trained to run the group sessions and modular courses GP trainers have to attend. Topics that are covered in these modular courses include GP-related expertise and teaching skills.
Teaching staff received training on how to conduct practice visits and preside over the peer groups during two separate three-hour sessions held for this study. They were trained to use a standardized agenda in which time to be used was specified for each part of the program. We developed a standardized agenda and a format for the minutes of the practice visits and peer-group
meetings. In addition, the teaching staff was taught how to give (practice visits) or trigger (peer groups) feedback on how SMART action plans were defined. The teaching staff prepared the peer-group meeting or practice visit by studying the feedback overviews and PDPs of the participating GP trainers.

*Instruments*

A 5-point Likert scale was used to score each SMART-criterion. Anchors were defined for the highest (5) and lowest (1) scores. For the criterion Measurable, the score of three was also defined. In addition, questions were formulated to aid scoring each criterion. Subsequently, the instrument was tested by the first and last authors, both of whom are experienced GPs with long-term involvement in GP Specialty Training. They independently scored the goals of 10 randomly selected PDPs using the scoring instrument. If comparison of the scores showed differences greater than one point, or the authors could not decide on a score, the matter was discussed. If the questions and anchors appeared to be insufficient for capturing the differentiation in the material to be scored, then they were adjusted. The pilot made it clear that the action plans gave too little information on the criterion Acceptability (A). Therefore, this criterion was not included in the final instrument. The 10 action plans used in the pilot were scored again using the definitive instrument. In addition to scoring SMARTness, the use of the AART criteria was scored with a yes/no tick box. See appendix for the final version of the scoring instrument used.

*Analyses*

Dichotomous data are presented as percentages. Scores on a Likert scale are presented as averages. We used multivariate analyses for repeated measures to test the effect of the intervention (peer group or practice visit) on the improvement in scores (baseline compared to follow-up) on the SMART-criteria (specific, measurable, realistic, time-bound). In cases where the effects in this overall analysis were significant, post-hoc analyses were performed to test the effects per SMART-criterion.
Results

Response rate
We originally invited all 85 trainers who were active as trainers of 1st-year GP trainees. They were randomly assigned to either the practice-visit group or the peer group (fig.1). Two were excluded because of their conflicting roles as both trainers and teaching staff members of the GP training institute at the time of our study. Ten others could not participate due to retirement before the end of the study, not being available for the whole study period or logistical problems. 25 trainers from the practice-visit group and 32 trainers from the peer group submitted baseline and follow-up PDPs. These are included in the analyses. GP trainers who did not submit a PDP gave restricted time as their reason for not doing so.

Personal Characteristics
The average age of trainers in years was 51 in the practice-visit group and 49 in the peer group (table 1). Both groups consisted of more male than female trainers (peer group 63%; practice-visit group 68%). The number of part-time trainers in the peer group was considerably higher than in the practice-visit group (78% versus 52%).

Table 1. Personal characteristics of GP trainers and their practices

<table>
<thead>
<tr>
<th></th>
<th>Practice visit (n=25)</th>
<th>Peer group (n=32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>32.0</td>
<td>37.5</td>
</tr>
<tr>
<td>Male</td>
<td>68.0</td>
<td>62.5</td>
</tr>
<tr>
<td>Average age</td>
<td>51</td>
<td>49</td>
</tr>
<tr>
<td>Part-time GP trainer (%)</td>
<td>52.0</td>
<td>78.1</td>
</tr>
<tr>
<td>Practice Size (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-handed</td>
<td>24.0</td>
<td>18.8</td>
</tr>
<tr>
<td>Small (up to three doctors)</td>
<td>56.0</td>
<td>56.3</td>
</tr>
<tr>
<td>Health care centres (more than 3 doctors)</td>
<td>20.0</td>
<td>25.0</td>
</tr>
<tr>
<td>Practice Setting (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City</td>
<td>36.0</td>
<td>53.1</td>
</tr>
<tr>
<td>Urban</td>
<td>48.0</td>
<td>28.1</td>
</tr>
<tr>
<td>Rural</td>
<td>16.0</td>
<td>18.8</td>
</tr>
<tr>
<td>In 3rd or 4th year of curriculum for starting GPs</td>
<td>40.0</td>
<td>50.0</td>
</tr>
<tr>
<td>Working in appraised practice</td>
<td>28.0</td>
<td>40.6</td>
</tr>
</tbody>
</table>

5 The Dutch Society of General Practitioners (NHG) appraises practices. In this (voluntary) appraisal the SMART-criteria are used in defining practice-based goals. GP trainers who work in an appraised practice therefore have experience of making goals SMART.
Figure 1. Flow chart of study group

**GP trainers**
- 38 trainers randomised and instructed to write PDP
- 32 trainers followed up
- 28 trainers had baseline PDP
- 8 trainers had no baseline PDP
- 32 trainers had baseline PDP and follow-up
- 4 trainers had no follow-up

**Non-participation**
- 2 trainers retired
- 1 trainer had conflicting roles
- 4 trainers had logistic problems

**Participation**
- 37 trainers participated in PV
- 8 trainers were sick
- 2 trainers were retired
- 1 trainer had conflicting roles
- 4 trainers had logistic problems

**Technical problems**
- 35 trainers participated in PG
- 0 trainers had baseline PDP
- 0 trainers had no baseline PDP
- 29 trainers had baseline PDP and follow-up
- 4 trainers had no follow-up
- 20 trainers had baseline PDP and follow-up
- 4 trainers had no follow-up

**Follow-up and analysis**
- 29 trainers followed up and analysed
- 4 trainers followed up and analysed
- 3 trainers followed up and analysed
- 22 trainers followed up and analysed
Tailoring to educational needs

Smartness scores
The majority of GP trainers (>80%) defined their goals using the subheadings of AART (aim, action, result and time span). This was the case in both the practice-visit group and peer group, although the baseline scores were slightly higher in the latter. Use of the subheadings increased in both groups after the intervention, except for Result in the practice-visit group. Time Span in the peer group decreased from 100% at baseline to 96.9.

The SMARTness scores (specific, measurable, realistic and time-bound) improved from baseline to follow-up irrespective of the intervention (F (1,56)=56,41 p<0.000).

Although not all SMART-criteria benefitted equally from the intervention (F (1,54)=6,50 p<0,001) post-hoc analyses revealed that the improvement in scores were significant for all SMART-criteria: Specific (F(1,56)=69,70 P<0,00), Measurable (F(1,56)=21,97 P<0,00), Realistic (F(1,56) =39,8 P<0.00) and Time-bound (F(1,56)=19,74 P<0.00). Overall, scores in the peer group were higher than in the practice-visit group (F (1,56)=5,4 P<0,024).

Discussion

Discussing the Smartness of the action plans for goals that are defined in a PDP improves this same SMARTness. This is congruent with literature about the positive effects of training the goal-setting abilities of case managers working in revalidation and mental-healthcare.14,15

Our study also shows that this improvement occurs after both a practice visit and participation in a peer group - improvement in scores was seen for both modes of intervention, with no significant differences between the two. This is encouraging for any educational setting addressing the quality and development of its clinical teachers since resources and time are usually limited everywhere and therefore have to be used as efficiently as possible. The cost-effectiveness of peer groups may well prompt a change towards the increased use of peers in appraisal, but there are further potential benefits. In peer groups, participants benefit from discussing the feedback received by other participants, too, and in addition, feedback from peers seems to be highly appreciated and leads to better practice outcomes than non-physician observers.12,13
### Table 2. Percentage of PDPs in which subheadings of AART were used

<table>
<thead>
<tr>
<th>Practice visit</th>
<th>Peer group</th>
<th>Baseline (n=25)</th>
<th>Follow-up (n=25)</th>
<th>∆ baseline follow-up</th>
<th>Baseline (n=32)</th>
<th>Follow-up (n=32)</th>
<th>∆ baseline follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aim mentioned (% of all)</td>
<td>84 (9.7%)</td>
<td>100 (0.0%)</td>
<td>69 (3.8%)</td>
<td>3.8</td>
<td>100 (0.0%)</td>
<td>100 (0.0%)</td>
<td>6.2</td>
</tr>
<tr>
<td>Action mentioned (% of all)</td>
<td>84 (9.7%)</td>
<td>92 (0.4%)</td>
<td>69 (3.8%)</td>
<td>6.9</td>
<td>100 (0.0%)</td>
<td>100 (0.0%)</td>
<td>3.1</td>
</tr>
<tr>
<td>Result mentioned (% of all)</td>
<td>80 (1.6%)</td>
<td>80 (1.6%)</td>
<td>90.6 (0.8%)</td>
<td>9.6</td>
<td>80 (1.6%)</td>
<td>100 (0.0%)</td>
<td>3.4</td>
</tr>
<tr>
<td>Time span mentioned (% of all)</td>
<td>80 (1.6%)</td>
<td>100 (0.0%)</td>
<td>96 (0.8%)</td>
<td>9.6</td>
<td>80 (1.6%)</td>
<td>100 (0.0%)</td>
<td>3.4</td>
</tr>
</tbody>
</table>

### Table 3. Mean Smartness scores (SD)

<table>
<thead>
<tr>
<th>Practice visit</th>
<th>Peer group</th>
<th>Baseline (n=25)</th>
<th>Follow-up (n=25)</th>
<th>∆ baseline follow-up</th>
<th>Baseline (n=32)</th>
<th>Follow-up (n=32)</th>
<th>∆ baseline follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Specific (SD)</td>
<td>1.72 (0.45)</td>
<td>3.16 (0.80)</td>
<td>1.44 (2.75)</td>
<td>2.75 (0.88)</td>
<td>3.66 (0.83)</td>
<td>0.91 (3.1)</td>
<td></td>
</tr>
<tr>
<td>Mean Measurable (SD)</td>
<td>2.28 (0.84)</td>
<td>2.72 (1.21)</td>
<td>0.44 (2.97)</td>
<td>1.21 (0.93)</td>
<td>3.53 (0.88)</td>
<td>0.56 (3.1)</td>
<td></td>
</tr>
<tr>
<td>Mean Realistic (SD)</td>
<td>2.32 (0.96)</td>
<td>3.32 (0.81)</td>
<td>1.00 (3.03)</td>
<td>1.00 (0.78)</td>
<td>3.66 (0.90)</td>
<td>0.63 (3.1)</td>
<td></td>
</tr>
<tr>
<td>Mean Time-bound (SD)</td>
<td>2.64 (1.04)</td>
<td>3.64 (1.04)</td>
<td>1.00 (3.64)</td>
<td>1.00 (0.88)</td>
<td>3.75 (0.88)</td>
<td>0.37 (3.1)</td>
<td></td>
</tr>
</tbody>
</table>
There were also a number of remarkable results. The baseline SMARTness scores were fairly high. We could argue that this was because we instructed trainers to use the AART format, which is also used by GP trainees in their portfolios and as a consequence, GP trainers are already familiar with it. Surprisingly and despite randomization, baseline SMARTness scores were higher in the peer group than in the practice-visit group. Perhaps GP trainers pay more attention to their action plans in anticipation of feedback from peers as compared to feedback during a practice visit from a single teaching staff member. Another explanation may be that in the peer group more GP trainers (50% vs. 40%) had already participated in the voluntary GP practice appraisal that is conducted by the Dutch Society of General Practitioners (NHG). For this appraisal, practice improvement plans are required for which the SMART-criteria are used, too. Interestingly the percentage of participants who submitted their baseline PDP was higher in the PG (68% versus 89%). Also the baseline scores were higher. It could be that the anticipation of meeting peers stimulates GP trainers to prepare themselves better. However this stimulating effect does not persist in finalizing their follow-up PDP’s after the peer groups (90 versus 91 %).

Another interesting finding is that the SMART-criterion Specificity seemed to benefit most from the interventions. Based on our observations during this study, we hypothesize that this is the criterion on which most feedback was given both in the peer groups and during the practice visits. We find it striking that it appears to be possible to define SMART-criteria by setting anchors per criterion in such a way that it is possible to measure change in SMARTness. As far as we know, there are no previous publications on this subject.

**Limitations**

There were a number of limitations to our study. First, much effort was required to get GP trainers to submit their PDPs. Since it was part of a research project, however, we were able to invest considerably in compliance. Doing so in a non-research setting may well put too much strain upon the resources available. On the other hand, going through the PDP cycle several times will mean that GP trainers get used to it – experience it as less time consuming – and eventually may appreciate it more.

Another limitation is that we did not use a control group. We can therefore not be sure whether the intervention caused the improvement of the SMARTness
of action plans, or merely the fact that GP trainers revised their goals. The small number of participants is a limitation, too. Unfortunately, this burdens many research projects in medical education. Multi-centre trials are needed to achieve larger numbers of participants, but this inevitably leads to difficulties in guaranteeing similar conditions at all participating centres.

Future research
Our presumption that ‘the Smarter an action plan is defined, the likelier its successful realization will be’ has not been tested by us, or, to our knowledge, by anyone else. It is the conclusion of a logical train of thought but if proven wrong, it would shake up the whole idea of the effectiveness of using the SMART-criteria, or, for that matter, any other mode of intervention to improve goal-setting abilities. So it remains an assumption, not a fact. In future research it would be a challenge to find evidence in its favour. In this study we looked at the quality of action plans from the methodological point of view: Are they defined SMARTly? In a previous article we described the relationship between the content of the goals and written competency-based feedback received by GP trainers\textsuperscript{11}. We have not investigated the quality of the content of the plans. It would be interesting to investigate the influence of peer groups versus practice visits regarding the quality of the content of the action plans.
References