*Counselling and Psychotherapy Research*

WILEY

ORIGINAL ARTICLE OPEN ACCESS

# Are We Practicing Deliberately? — Assessment of Deliberate Practice in Psychotherapy

Stefan Blümel[1] | Charlotte Schmidt-Pleines[2] | Magdalena Pape[1] | Sabine Steins-Loeber[1]

[1]Department of Clinical Psychology and Psychotherapy, Otto-Friedrich University of Bamberg, Bamberg, Germany | [2]Department of Developmental Psychology and Psychological Diagnostics, University of Koblenz, Koblenz, Germany

**Correspondence:** Stefan Blümel (stefan.bluemel@uni-bamberg.de)

**Received:** 13 November 2025 | **Revised:** 8 June 2026 | **Accepted:** 12 June 2026

**Keywords:** assessment | deliberate practice | psychotherapy | psychotherapy training | therapist effects

## ABSTRACT

**Objective:** Deliberate Practice (DP) is a promising approach to improve psychotherapy outcomes and skill acquisition. Based on theoretically derived core components of DP in psychotherapy, the aim of this study is to develop assessment instruments for measuring behaviours related to and attitudes towards DP.

**Method:** Two item sets were generated associated to DP's core components in psychotherapy and characteristics of DP. Two exploratory factor analyses were conducted to develop two scales (DP Behaviours [DPB] and DP Attitudes [DPA]). Data were collected from prospective psychotherapists in Germany ($N = 199$). Furthermore, in a subsample ($N = 138$) construct validity was tested using correlational analysis with convergent and discriminant constructs.

**Results:** The analysis supported a two-factor structure (*activities* and *feedback and expert involvement*) for and construct validity of the DPB scale, indicating associations in expected directions. For the DPA scale, one factor emerged indicating a generalised attitude towards excellence and improvement. Construct validity was partially supported, with convergent correlations in expected directions, but inconsistencies regarding discriminant validity.

**Conclusion:** The scales developed in this study demonstrate potential for linking self-assessments of DP to psychotherapy outcomes. Future research needs to consider the predictive validity of DP's core components in relation to psychotherapy outcomes and broader factors influencing DPB and DPA, such as a culture of excellence or trusting and healthy relationships with experts.

## 1 | Introduction

### 1.1 | Deliberate Practice

Deliberate practice (DP) is defined as 'the individualized training activities specially designed by a coach or teacher to improve specific aspects of an individual's performance through repetition and successive refinement' (Ericsson and Lehmann 1996, 278–279).

In psychotherapy research, DP is viewed as a promising variable, that could partially explain differences in the effectiveness of therapists (Wampold and Owen 2021). Therefore, it is suggested to improve therapeutic outcomes (Tracey et al. 2015). This could be achieved via training and mastering psychotherapeutic skills associated with outcomes (e.g., facilitative interpersonal skills; Anderson et al. 2016). Concerning the direct relationship of DP with outcomes, evidence is mixed but promising (Chow et al. 2015; Goldberg, Babins-Wagner,

**Implications for Practice**

- The newly developed Deliberate Practice Behaviours scale (DPB) and Deliberate Practice Attitudes scale (DPA) are a promising method to measure several facets of Deliberate Practice.
- DPB and DPA could inform therapists which components of DP need more or less focus during training.
- DPB and DPA also have the potential to inform Deliberate Practice implementation processes and help educate on Deliberate Practice components and roadblocks.

**Implication for Policy**

- DP research needs to be more comparable, which the DPA and DPB scales have the potential to enable. Drawing on strong research, guardrails for training and continued education can be developed implementing deliberate practice to improve overall expertise and therapeutic outcomes.

et al. 2016; Janse et al. 2023). Debates have emerged around DP in general, noting its changing definition complicates research and operationalisation and therefore also the measurement and comparability of findings (Hambrick et al. 2021). Similarly, different understandings concerning DP in psychotherapy exist (Clements-Hickman and Harris 2024; Diamond et al. 2025; Mahon 2023).

## 1.2 | DP Models in Psychotherapy

A prominent approach to DP is put forward by Scott Miller, Daryl Chow and colleagues. Miller et al. (2020) include setting of individualised learning goals based on psychotherapy outcomes, immediate feedback, a coach and successive refinement to improve specific skills as an interactive system. Feedback in this model means, to get a response to your performance and learning process (e.g., by clients). This translates to using standardised measures for establishing an outcome baseline and build on it by identifying areas of improvement (Miller et al. 2020, 43–47). Another approach is put forward by the working group around Tony Rousmaniere and Alex Vaz. Rousmaniere (2017, 2024) underlines that DP involves repeated practice of specific skills with expert feedback, the observation of ones work, setting of learning goals just beyond the performer's ability and assessment of performance. Rousmaniere (2024, 96) illustrates DP as a circular system based on the above mentioned processes and continued throughout ones career. Emphasis in this model lies on the development of predetermined micro-skills as a basis of training, solitary practice due to systemic limitations such as scarcity of DP supervisors and the importance of simulation-based training (Rousmaniere 2024, Chapter 9).

However, these models differ in structure, emphasis and operationalisation. Rousmaniere (2024) explains that there exist two options to determine the exercises to practice (assignment by coach or picked by the learner based on own judgement) while Miller et al. (2020) would emphasise the development of a learning project in cooperation with a coach that can be traced back to a common factor oriented domain. Additionally, a pronounced difference exists in the targets of practice. It is emphasised to only practice in domains of improvement that are strongly associated with client outcomes (Miller et al. 2020). On the other hand it is pointed out that therapy model specific skills can be targets of DP as well (e.g., Boswell and Constantino 2022; Levenson et al. 2023). Vaz and Rousmaniere (2022) argue that 'DP research only informs us on *how* to practice, not on *what* to practice'. (13). This is underlined by the contextual model of psychotherapy highlighting the importance of the therapists ability of convincingly and competently delivering a rationale and treatment to a client (Wampold and Imel 2015). Another difference persists concerning the information of the practice process. For Miller et al. (2020) the practice needs to be rooted in averaged client data to identify coherent patterns for improvement. Rousmaniere (2024, 101) argues to add other sources (e.g., client verbal report, coach judgement, qualitative data and collateral information). Consequently, DP begins informed by an individual, specific case the therapist is challenged by. Concerning the whole of the DP process, the sequence and interplay of goal identification and repetition and successive refinement are explained differently in the two models. For an excellent visual representation of both DP models see Diamond et al. (2025, 4).

This overall heterogeneity has not gone unnoticed and calls for adjustments, unification and coherent operationalisation of DP models in psychotherapy (Clements-Hickman and Harris 2024; Diamond et al. 2025; Mahon 2023; Nurse et al. 2024). Drawing on the aforementioned models of DP in psychotherapy, Diamond et al. (2025) suggest four core components to be a common ground and applications of DP in psychotherapy should incorporate them: setting learning objectives just beyond the learners' current ability targeted at a specific deficit, behavioural repetition and successive refinement of specific skills or factors leading to performance improvement, the involvement of an experienced coach or supervisor and the incorporation of immediate and continuous feedback. The authors of this article argue another component is present in both DP models that could be called: structuring, self-confrontation and ongoing evaluation of the professional development. Miller et al. (2023) lay out a complex, detailed and interactive approach to this component in their 'Field guide to better results', while Rousmaniere (2024, 103) describes the necessity of continuous and long-term assessments of DPs' effectiveness in relation to client outcomes.

## 1.3 | Measuring DP

Due to the mentioned inconsistencies in DP research in psychotherapy, operationalisation of DP is heterogeneous. Previous research has focused on measuring outcome and domains of improvement (Miller et al. 2020). So far, studies investigating the DP-outcome relationship used a questionnaire, estimating time allocated at some of DP's components (RAPID). It was identified as potential for improvement (Chow et al. 2015; Janse et al. 2023). Other measurement approaches exist, using medical instruments. Unfortunately, none are

adapted, psychometrically analysed and validated for a psychotherapy context (Duvivier et al. 2011; Lipp 2019; Moulaert et al. 2004). An alternative way to measure DP is through diaries that track activity type, duration and difficulty (Miller et al. 2020; Rousmaniere 2017). However, diary methods are time-consuming and add complexity, leading to calls for feasible and theory-based instruments (Janse et al. 2023; Vaz and Rousmaniere 2022). Hence, until today, no measure exists assessing to what extent DP was used. Additionally, no instrument exists that assesses enabling factors of DP. This is surprising, as environmental and intrinsic factors are assumed to play a crucial role in the development of expertise through DP (Ericsson et al. 2009; Ericsson and Pool 2016). Furthermore, research on therapist effects indicates that certain attitudes (e.g., expectations for long-term success, attitudes towards feedback or professional self-doubt) are associated with improved outcomes (Wampold and Owen 2021). Mahon (2024) found that beliefs about positive potential influences of DP or the adverse effects of not using DP (stagnation or even deterioration as mentioned in Goldberg, Rousmaniere, et al. 2016) were one of the keys for psychotherapists to engage in DP. Fishbein and Ajzen (2009) suggest that attitudes or beliefs result in an intention to and behaving a certain way. This implies that attitudes towards DP could enable the execution of DP behaviour, which in turn possibly influences therapeutic outcomes.

In sum, existing measures focus on 'What?' is practiced than on 'How?' or 'How much in what way?'. Especially the latter ones are of relevance for the classification of a practice to be DP (Husby 2025), which in turn allows for an assertion to what degree DP is associated with outcomes. To conclude, it seems warranted to inform future DP research with measures that operationalise facets of DP such as behaviour or attitudes. These operationalisations should reflect the core components of DP and be distinguishable from related constructs. Unlike broader constructs such as general professional development (Bennett-Levy 2019) or achievement motivation (McClelland et al. 2012), DP is not primarily defined by a general wish to improve or by valuing professional growth, but rather DP refers to a specific mode of practice characterised by structured, effortful, feedback-informed practice directed at improving clearly defined aspects of therapeutic performance.

### 1.4 | The Present Study

Current research implies that interpersonal skills (e.g., facilitative interpersonal skills) are associated with improved therapeutic outcomes as well as practicing deliberately. Furthermore, DP can be used to improve therapist skills which could point to an indirect effect on therapeutic outcome. Heterogeneous approaches to DP in psychotherapy, and consequently, the operationalisation and measurement of DP in the field of psychotherapy are evident though. Therefore, the present study aims for the development of a new measurement approach to DP in the acquisition of psychotherapeutic skills. This will be achieved by developing two instruments to assess behavioural and motivational aspects of DP. The motivational domain is included as it assesses the internal conditions of therapists that make sustained DP more likely. However, as practice to qualify as DP, it is also necessary to measure how therapists actually practice, which will be addressed by the behavioural domain. To summarise, DP cannot be adequately assessed by motivation or practice behaviour alone, because sustained skill development requires both the willingness to engage in challenging practice and the actual implementation of DP-consistent activities. For both instruments, exploratory factor analysis will be conducted, and construct validity will be investigated.

## 2 | Method

### 2.1 | Participants

Two hundred and eleven participants were recruited. Twelve participants were excluded due to careless responding indicated by three items testing comprehension of the scale instructions. The final sample consisted of $N = 199$ participants. Of those, $n = 138$ participants also filled in data for construct validity and $n = 61$ participants took also part in an ongoing evaluation study examining the development of psychotherapeutic competencies (data on the evaluation are not reported here). This overall sample of prospective psychotherapists consists of advanced psychotherapy students ($n_{\text{undergraduate students}} = 37$, $n_{\text{graduate students}} = 87$) having practical experience in treating clients and psychotherapy trainees ($n_{\text{trainees}} = 75$) currently working in psychiatric or in outpatient clinics. Our sample is remarkably affected by the implementation of DP in training contexts, which makes them particularly appropriate for our aims. This includes DP-based courses for skill acquisition during psychotherapy master's programs and the provision of training facilities and materials (e.g., DP manuals, DP instructions) for voluntary DP group practice outside of the course hours. Participants were on average 27.3 years old (SD = 6.08, range 18–49) and mostly female ($n_{\text{female}} = 169$, $n_{\text{male}} = 29$, $n_{\text{diverse}} = 1$).

### 2.2 | Sample Size

Research by MacCallum et al. (1999) suggests that communalities of investigated items and determination of extracted factors are indicators for the required sample size to proceed with factor analysis. Therefore, a sample size of approximately 200 participants is targeted due to simulations showing that for most cases of communalities and factor determinations a sample of 200 participants is sufficient to produce stable factor solutions. Moreover, a sample size of 200 would be within the typical range of the several rules of thumb when calculating with 34 items (Eid et al. 2017).

### 2.3 | Procedure

Participants were recruited through social media, mailing lists from psychotherapy departments and training institutes. Data collection lasted from May 2023 to July 2024. To incentivise participation, ten vouchers worth 20 Euros were randomly raffled among participants completing data for construct validity. The study was performed in line with the Declaration of Helsinki. In accordance with local regulations, no ethical approval was required for the survey data collection. Informed consent was obtained from all participants.

## 2.4 | Scale Development

As a first step, two item pools were generated in German by the first two authors with the first scale containing items targeting specific DP-related behaviour in relation to skill acquisition and the second scale consisting of items describing attitudes towards characteristics of DP. Second, both item sets were reviewed for plausibility and phrasing by three faculty colleagues who have been incorporating DP in their teaching and are simultaneously psychotherapists familiar with applying DP in their own therapeutic practice. As a last step, based on the feedback, items were revised and underwent a final evaluation for content validity in the context of psychotherapy and training as well as for comprehensibility. This procedure followed recommendations for item generation (Moosbrugger and Kelava 2020).

The first item set aggregated 34 items targeting DP behaviours (e.g., 'I used the feedback from experts to adjust my learning objectives'. or 'I repeated the exercises on specific therapeutic skills several times'.). Items could all be allocated to one of DPs core components (Diamond et al. 2025). Therefore, the DP behaviours scale consisted of items describing specific activities that comprise DP such as goal-setting, repeated, behavioural rehearsal, structuring the learning process and feedback seeking regarding improvements and adaptation of the beforementioned components. The second item set consisted of 15 items targeting attitudes towards DP (e.g., 'It is important for me, to set myself challenging goals'. or 'I think that repetitive practicing of therapeutic skills makes sense'.). To summarise, the DP attitudes scale comprised of items assessing a motivational stance towards the characteristics of DP (e.g., striving for excellence, openness to improve through feedback, training at the edge of current capabilities). Items were answered on a five-point Likert-type scale with the values on the behaviour scale ranging from 'never' to 'always' and for the attitudes scale ranging from 'totally disagree' to 'fully agree'. All items included can be accessed in the Supporting Information.

## 2.5 | Assessment of Construct Validity

### 2.5.1 | Beharrlichkeit Und Beständiges Interesse (BISS-8)

The BISS-8 (Schmidt et al. 2019) is a validated German version of the Grit-scale (Duckworth et al. 2007) which measures perseverance and passion for long-term goals and is considered to be related to DP (Duckworth et al. 2011; Mahon 2024; Rousmaniere 2017). The items are rated on a 5-point Likert-type scale ranging from 1 = not at all like me to 5 = very much like me. Internal consistency in the present sample for the overall Grit score was good with $\alpha = 0.80$.

### 2.5.2 | Leistungsmotivationsinventar-Kurzversion (LMI-K)

The LMI-K (Schuler and Prochaska 2001) is a 30-item short version of an inventory designed to measure orientation towards general professional practice. The underlying construct of need for achievement could be described as a drive to reach mastery or important accomplishments (Finogenow 2017), therefore, being a related construct to DP. All items were rated on a 7-point Likert-type scale ranging from 1 = does not apply at all to 7 = applies in full. Internal consistency in the present sample for the LMI-K score was excellent with $\alpha = 0.93$.

### 2.5.3 | Big Five Inventory (BFI-2-S)

The German short version of the Big Five Inventory-2 by Rammstedt et al. (2020) consists of 30 items with internal consistencies for the five subscales between $\alpha = 0.72$ and 0.78 in the present sample. Items are rated on a 5-point Likert scale ranging from 1 = completely disagree to 5 = fully agree. There is no evidence of Big Five being associated with DP apart from conscientiousness, which is associated with Grit and therefore should be assumed for convergent validity (Duckworth et al. 2007). Consequently, the remaining Big Five subscales should not consistently correlate with the DP scales and can therefore be utilised to assess discriminant validity.

## 2.6 | Data Analysis

For analyses Jamovi (The Jamovi Project 2024) and IBM SPSS Statistics for Windows, Version 29.0.2.0 (IBM Corp 2023) were used. Alpha levels were set at 5%. Mean values were imputed for missing data under 5% (DP Behaviours and measures for construct validity), and multiple imputation was used for missing values between 5% and 10% (DP Attitudes) (Tabachnick and Fidell 2007). Item selection was based on several criteria: item-total correlation, change in Cronbachs alpha, item difficulty, loadings and content wise importance and relevance. Before proceeding with exploratory factor analysis, a measure of sampling adequacy (KMO) and a Bartlett test were calculated. KMO values should be at least 0.50 (Field 2013). For factor extraction we resorted to principle axis factoring (PFA) (Eid et al. 2017). A combination of criteria was used to decide on the number of factors (Moosbrugger and Kelava 2020). Those were the Kaiser-Guttman criterion, scree plot, parallel analysis and the content of the items, which are combined to a factor. For easier interpretation of underlying factors, direct oblimin rotation was used, allowing factors to correlate (Eid et al. 2017).

# 3 | Results

## 3.1 | DP Behaviours Scale (DPB)

Item-total correlation revealed seven items that did not reach sufficient item-test correlations ($r_{it} < 0.3$). Cronbachs alpha marginally improved the scale if either one of the seven items was discarded. MSA values were below the cut-off for three of the seven items. After reviewing the content of the respective items and identification of sufficient alternatives in the remaining item pool, all items failing to achieve one of the selection criteria were excluded. Item difficulties followed a normal distribution. Overall KMO (MSA = 0.91) and the Bartlett test ($\chi^2(351) = 2495$,

$p < 0.001$) allowed to continue with PFA. Item communalities ranged from 0.12 to 0.78. Two criteria pointed towards a two-factor solution (Kaiser-Guttmann criterion, scree plot). Parallel analysis favoured a three-factor solution (see Figure 1). A two-factor solution is favoured. Items were then oblimin rotated. All items loaded on either one of the factors and all loadings were above 0.3 (see Supporting Information). The solution accounted for 41.2% of the variance. The first factor explains 25.1% and the second 16.1%. The factors showed a medium to high correlation ($r = 0.43$). Internal consistency of the subscale adapted from the first factor was excellent ($\alpha = 0.90$) and high for the second ($\alpha = 0.87$). The first factor comprises 18 items depicting different activities associated with DP such as setting of learning objectives, repeated practice or confrontation with one's own practice. Therefore, the according subscale is labelled *activities*. The second factor is composed of nine items describing the involvement of a coach and feedback towards the activities. The subscale is labelled *feedback and coach involvement*.

### 3.2 | DP Attitudes Scale (DPA)

Item-total correlations revealed one item with lower associations than $r_{it} < 0.3$ and excluding the item improved Cronbachs alpha. Another item was almost below the cut-off value ($r_{it} = 0.32$) and dropping the item would result in an improved Cronbachs alpha. Visual inspection of the item difficulties revealed ceiling effects. Though, all item difficulties are below $P \leq 95$. In sum, two items were excluded. Overall KMO (KMO = 0.83) and Bartlett test ($\chi^2(78) = 707$, $p < 0.001$) allowed to continue with PFA. Item communalities ranged between 0.14 and 0.50. After conducting PFA, the Kaiser-Guttmann criterion supported a one-factor solution. Scree plot favoured a two-factor structure, although interpretation of the 'elbow' was difficult and the 'steepness' of the drop after the first factor was small. Parallel analysis suggested four factors (see Figure 2). After reviewing the content of the items associated with each factor solution, no meaningful factor description was discernable. Therefore, a one-factor solution

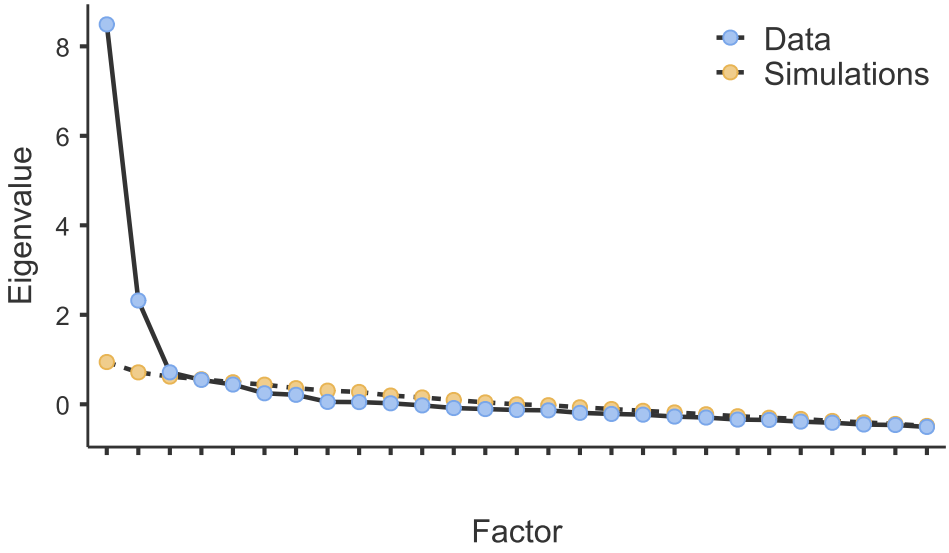


**FIGURE 1** | Scree plot and parallel analysis for DPB.

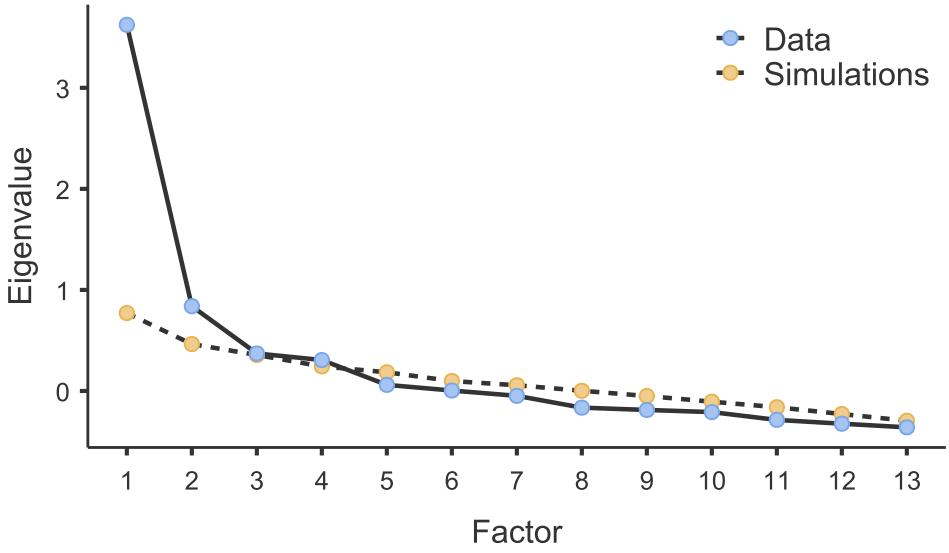


**FIGURE 2** | Scree plot and parallel analysis for DPA.

was retained. The factor accounts for 27.9% of variance. Internal consistency for the final scale was high ($\alpha = 0.82$).

### 3.3 | Convergent Validity

Both newly developed scales seem to be associated with one another in a meaningful pattern that fits the content of the factors. The DPB subscale *activities* and the DPA are significantly associated ($r = 0.22$, $p < 0.01$) while the *feedback and coach involvement* subscale and DPA aren't ($r = 0.1$, $p > 0.05$), which seems reasonable due to the focus of the DPA relying on attitudes towards practice, excellence in performance and continuous improvement rather than receipt of feedback and engaging with coaches.

The DPB subscale *feedback and coach involvement* was not associated with LMI-K ($r = 0.16$, $p > 0.05$) and less strongly associated with Grit ($r = 0.20$, $p < 0.05$). This could be due to the items being focused at involving another person while Grit and LMI-K are rather focused on the individual (Table 1). The remaining results further support convergent validity for DPA and DPB due to correlations with convergent measures being in the expected directions and mostly medium to high in size (see Table 1).

### 3.4 | Discriminant Validity

#### 3.4.1 | DPB

For discriminant validity, the BFI-2-S (Rammstedt et al. 2020) was used (see Table 2). As expected, there was a significant positive medium correlation between conscientiousness and the *activities* subscale, but not for the *feedback and coach involvement* subscale, emphasising the association between activities such as setting of learning objectives or practicing with a conscientious trait. For the other dimensions, no coherent at least medium association could be detected, further strengthening construct validity of the DPB scale.

**TABLE 1** | Correlation matrix for convergent validity between DPB subscales, DPA, LMI-K and BISS-8/Grit.

| | DPA | DPB (activities) | DPB (feedback and coach involvement) |
|---|---|---|---|
| LMI-K | 0.33*** | 0.29*** | 0.16 |
| BISS-8/GRIT | 0.25*** | 0.43*** | 0.20* |

*Note:* $N = 138$.
*$p < 0.05$.
***$p < 0.001$.

#### 3.4.2 | DPA

Results for the DPA were mixed. In line with expectations, the DPA exhibits a significant medium association with conscientiousness ($r = 0.29$, $p < 0.001$). Contrary to expectations, there was also a relation between agreeableness and the DPA ($r = 0.41$, $p < 0.001$). Furthermore, DPA was associated negatively with neuroticism ($r = -0.19$, $p < 0.05$) and positively with openness ($r = 0.25$, $p < 0.01$). Concluding, the results partly indicate discriminant validity.

Especially the DPB showed expected patterns, and it can be concluded that overall construct validity seems to be supported by the available data. For the DPA, results need to be treated with more caution. While results support convergent validity, data on discriminant validity is mixed.

## 4 | Discussion

DP potentially influences therapeutic outcomes through a direct or indirect (via skills) association. The operationalisation and the measurement of DP in psychotherapy have been heterogeneous though. Therefore, the present study aimed for a valid measurement of DP by developing two instruments that assess behaviours related to and attitudes towards DP in the domain of psychotherapy as long-term skill improvement requires not only motivation to engage in challenging practice, but also the actual execution of DP-related activities. Items relate to DP's core components and characteristics (Diamond et al. 2025). Initial evidence from exploratory factor analysis revealed promising results for both scales. The DPB consists of two factors which form the subscales *activities* and *feedback and coach involvement*. The first describes activities such as goal-setting or repeated behavioural practice whereas the second one comprises items that describe gathering of feedback from coaches or the collaborative adaptation of the DP activities with coaches. The DPA consists of one factor with items describing a motivational stance towards the demanding and ambitious characteristics of DP.

**TABLE 2** | Correlation matrix for discriminant validity between DPB subscales, DPA and BFI-2-S scales.

| | Openness | Neuroticism | Conscientiousness | Agreeableness | Extraversion |
|---|---|---|---|---|---|
| DPB (activities) | 0.13 | −0.13 | 0.32*** | 0.21* | 0.13 |
| DPB (feedback and coach involvement) | 0.15 | −0.14 | 0.13 | 0.07 | 0.19* |
| DPA | 0.25*** | −0.19 * | 0.29*** | 0.41*** | 0.14 |

*Note:* $N = 138$.
*$p < 0.05$.
***$p < 0.001$.

### 4.1 | DPB

Findings from the DPB could be aligned with initial concepts by Ericsson and Lehmann (1996, 278) defining DP as '[...] the individualized training activities especially designed by a coach or teacher to improve specific aspects of an individual's performance through repetition and successive refinement'. The subscale *feedback and coach involvement* match the definition and are complemented by the *activities* subscale. The latter one is more abstract due to different activities described through the allocated items. These items assess goal-setting activities, behavioural and repeated rehearsal of skills and the continued reflection and evaluation of the learning process. This evaluation is also addressed by items that capture continuous adjustment of goal-setting and behavioural repeated practice as well as the overall structure of practice. Interestingly, there did not emerge a partitioned structure in the *activities* scale allocating items to the different core components of DP. Feedback and coach items were clustered into one factor though. This could hint at the importance of feedback and the interconnecting characteristic of the coach who is highly involved in adapting DP activities. Feedback targeted exactly at areas of potential improvement such as the configuration of learning pathways guided by specific learning objectives just outside the practitioner's current ability or the joint adaptation of difficulties perceived by practitioners during repetitive, behavioural rehearsal could be informed by the DPB. This would be in line with findings by Westra et al. (2021) identifying individualization of feedback as a key factor during implementation of DP. Taken together, both DPB subscales could add value to current psychotherapy research. They enable the improvement of DP settings. First, to give an example, if the *activities* scale were rated less frequent, this could lead to more education on goal-setting techniques or how to incrementally change behavioural repeated practice exercises. On the other hand, low ratings on the *feedback and coach involvement* subscale could hint at possible improvements for coaches to be more approachable or to adapt their feedback to the practitioners' needs. Additionally, the measurement of both subscales could enable further insight into the mechanisms of skill acquisition and associations with psychotherapeutic outcomes by conducting component studies.

### 4.2 | DPA

As Ericsson et al. (2009) discussed, motivation seems to be one influential predictor that could differentiate the practitioners' degree of continued engagement in DP. Our results concerning discriminant validity were mixed. This effect could be attributed to the ceiling effects detected after item-difficulty inspection, which, in turn, could be caused by answers being influenced by social desirability that exhibits relations with conscientiousness and agreeableness (Ones et al. 1996). This indicates that our results could potentially be influenced by social desirability or in contrast to DPB, DPA could be associated with certain personality traits. In turn, this implies that DPB is potentially less susceptible to bias. Despite potential biases, the developed scale offers instructors a tool to promote DP implementation and targets for interventions concerning the rationale of DP. This is in line with calls for more education on the DP rationale before being integrated (Hill et al. 2020; McLeod 2022). Furthermore, as attitudes are important predictors of actual behaviour and the sustainability of the behaviour (Fishbein and Ajzen 2009), the assessment of DP attitudes could potentially inform psychotherapy research when examining and comparing the effectiveness of different learning contexts such as training facilities, universities, outpatient clinics or clinics.

### 4.3 | Limitations and Strengths

The present study underlies a couple of limitations and strengths that need to be considered. Our sample size was good for exploratory factor analysis drawing on MacCallum et al. (1999) findings, combining item communalities and determination of the found factors. We only assessed construct validity, so further research is needed to evaluate criterion and content validity. Although construct validity seemed to be promising for the DPB, there is still potential for improvement of the scales, especially for the DPA, with one option being to control for social desirability when administering the DPA. Furthermore, both scales underlie possible distortions due to being self-assessments, although especially convergent validity indicating associations with related constructs. Beyond that, our measures are not yet associated with observed psychotherapist behaviour or outcomes. On the other hand, the present study is, until today, the first opportunity to link psychometrically evaluated self-assessments of DP behaviour and attitudes to psychotherapy outcomes in a reliable and feasible way. Furthermore, findings of this study have theoretical support in the current psychotherapy research literature on DP.

### 4.4 | Future Directions

Evidently, both assessment instruments will need to undergo confirmatory factor analysis with a larger and broader sample size expanding to psychotherapists in practice. To further strengthen convergent validity for the DPB, diary methods need to be used to establish a link between the self-assessment and observed psychotherapist behaviour and then linking them to outcomes. Similar to component studies, it is warranted to investigate the predictive validity of DP's core components. The distinction found between the DPB subscales could be of interest for this endeavour. We further encourage researchers to translate the German version to other languages and replicate our findings. By measuring attitudes towards effortful practice, motivational differences between individuals leave the opportunity to develop interventions that enhance attitudes towards DP. However, DP should be viewed holistically, with the aim of developing a culture of excellence in which continuous improvement through feedback-informed, incremental and challenging practices is favoured. This approach should nurture strong and trusted relationships between peers and experts, enabling flourishing of intrapersonal skills such as self-compassion (Axelsson et al. 2024; Olsson et al. 2025) or responsiveness through emotion-regulation (Hatcher 2015; Orlowski et al. 2025). This would warrant future research to investigate the learning environment in training institutions, universities and outpatient clinics. This is also indicated by the proposal of the social environment being a highly influential enabler for DP (Ericsson and Pool 2016). This could be achieved using a holistic perspective

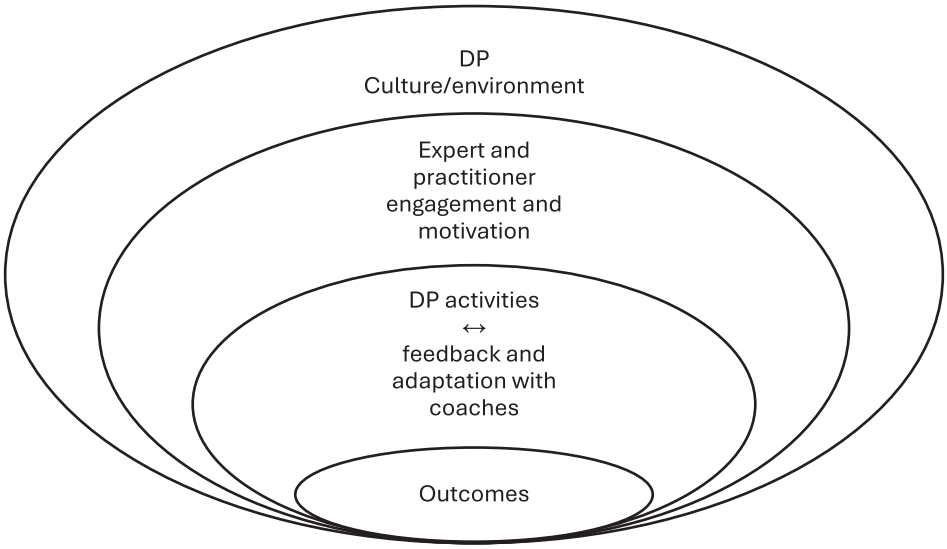


**FIGURE 3** | Holistic view of deliberate practice.

on DP that could be illustrated using layers of the learning environment with the superordinate goal of improving outcomes through DP (see Figure 3).

## 5 | Conclusion

Based on DP's core components described by psychotherapy research, two measures were developed targeting attitudes towards DP and specific DP-related behaviour. The findings for the DPB show that the items can be grouped into the *activities* and *feedback and coach involvement* subscales. The DPA seems to be unidimensional. This study also discusses the importance of taking a wider look at DP and the factors possibly influencing its successful implementation such as a positive attitude towards DP, which could be measured using DPA and a culture of excellence. We encourage researchers to apply these scales to inform studies in the future and to use them as a ground for comparison and collaboration on the quickly evolving topic of DP in psychotherapy. Furthermore, both scales can be used by practitioners to measure and monitor their own implementation of DP.

### Acknowledgements

The authors would like to thank Daniel Jäger, Frank Lörsch and Martin Lehe for their valuable support and feedback during theoretical discussions and item generation. Open Access funding enabled and organized by Projekt DEAL.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### References

Anderson, T., A. S. McClintock, L. Himawan, X. Song, and C. L. Patterson. 2016. "A Prospective Study of Therapist Facilitative Interpersonal Skills as a Predictor of Treatment Outcome." *Journal of Consulting and Clinical Psychology* 84, no. 1: 57–66. https://doi.org/10.1037/ccp0000060.

Axelsson, S., S. Kihlberg, P. Davis, and M. B. T. Nyström. 2024. "Psychotherapy Students' Experiences of Supervisee-Centred Supervision Based on Deliberate Practice, Feedback-Informed Treatment and Self-Compassion." *Counselling and Psychotherapy Research* 24, no. 2: 719–733. https://doi.org/10.1002/capr.12708.

Bennett-Levy, J. 2019. "Why Therapists Should Walk the Talk: The Theoretical and Empirical Case for Personal Practice in Therapist Training and Professional Development." *Journal of Behavior Therapy and Experimental Psychiatry* 62: 133–145. https://doi.org/10.1016/j.jbtep.2018.08.004.

Boswell, J. F., and M. J. Constantino. 2022. *Deliberate Practice in Cognitive Behavioral Therapy*. American Psychological Association.

Chow, D., S. D. Miller, J. A. Seidel, R. T. Kane, J. A. Thornton, and W. P. Andrews. 2015. "The Role of Deliberate Practice in the Development of Highly Effective Psychotherapists." *Psychotherapy* 52, no. 3: 337–345. https://doi.org/10.1037/pst0000015.

Clements-Hickman, A. L., and K. R. Harris. 2024. "Deliberate Practice for Psychotherapy Skills: Recommendations and Implications Based on the State of the Science." *Professional Psychology: Research and Practice* 55, no. 6: 563–572. https://doi.org/10.1037/pro0000592.

Diamond, G., B. Wlodek, S. Arthey, and S. Parker. 2025. "A Systematic Review of Deliberate Practice in Psychotherapy: Definitions, Operationalization, and Preliminary Outcomes." *Psychotherapy* 62, no. 2: 113–131. https://doi.org/10.1037/pst0000569.

Duckworth, A. L., T. A. Kirby, E. Tsukayama, H. Berstein, and K. A. Ericsson. 2011. "Deliberate Practice Spells Success: Why Grittier Competitors Triumph at the National Spelling Bee." *Social Psychological and Personality Science* 2, no. 2: 174–181. https://doi.org/10.1177/1948550610385872.

Duckworth, A. L., C. Peterson, M. D. Matthews, and D. R. Kelly. 2007. "Grit: Perseverance and Passion for Long-Term Goals." *Journal of Personality and Social Psychology* 92, no. 6: 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087.

Duvivier, R. J., J. van Dalen, A. M. Muijtjens, V. R. Moulaert, C. P. van der Vleuten, and A. J. Scherpbier. 2011. "The Role of Deliberate Practice in the Acquisition of Clinical Skills." *BMC Medical Education* 11, no. 1: 101. https://doi.org/10.1186/1472-6920-11-101.

Eid, M., M. Gollwitzer, and M. Schmitt. 2017. *Statistik und Forschungsmethoden: Mit Online-Materialien*. 5th ed. Beltz.

Ericsson, K. A., and A. C. Lehmann. 1996. "EXPERT AND EXCEPTIONAL PERFORMANCE: Evidence of Maximal Adaptation to Task Constraints." *Annual Review of Psychology* 47, no. 1: 273–305. https://doi.org/10.1146/annurev.psych.47.1.273.

Ericsson, K. A., K. Nandagopal, and R. W. Roring. 2009. "Toward a Science of Exceptional Achievement." *Annals of the New York Academy of Sciences* 1172, no. 1: 199–217. https://doi.org/10.1196/annals.1393.001.

Ericsson, K. A., and R. Pool. 2016. *Peak Secrets From the New Science of Expertise*. Houghton Mifflin Harcourt.

Field, A. 2013. *Discovering Statistics Using IBM SPSS Statistics*. 4th ed. SAGE.

Finogenow, M. 2017. "Need for Achievement." In *Encyclopedia of Personality and Individual Differences*, edited by V. Zeigler-Hill and T. K. Shackelford, 1–4. Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_537-1.

Fishbein, M., and I. Ajzen. 2009. *Predicting and Changing Behavior: The Reasoned Action Approach*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/ub-bamberg/detail.action?docID=668501.

Goldberg, S. B., R. Babins-Wagner, T. Rousmaniere, et al. 2016. "Creating a Climate for Therapist Improvement: A Case Study of an Agency Focused on Outcomes and Deliberate Practice." *Psychotherapy* 53, no. 3: 367–375. https://doi.org/10.1037/pst0000060.

Goldberg, S. B., T. Rousmaniere, S. D. Miller, et al. 2016. "Do Psychotherapists Improve With Time and Experience? A Longitudinal Analysis of Outcomes in a Clinical Setting." *Journal of Counseling Psychology* 63, no. 1: 1–11. https://doi.org/10.1037/cou0000131.

Hambrick, D. Z., B. N. Macnamara, and F. L. Oswald. 2021. "Is the Deliberate Practice View Defensible? A Summary of the Empirical Evidence and Reply to K. Anders Ericsson [Presentation]." https://osf.io/5nz4b.

Hatcher, R. L. 2015. "Interpersonal Competencies: Responsiveness, Technique, and Training in Psychotherapy." *American Psychologist* 70, no. 8: 747–757. https://doi.org/10.1037/a0039803.

Hill, C. E., D. M. Kivlighan, T. Rousmaniere, D. M. Kivlighan, J. A. Gerstenblith, and J. W. Hillman. 2020. "Deliberate Practice for the Skill of Immediacy: A Multiple Case Study of Doctoral Student Therapists and Clients." *Psychotherapy* 57, no. 4: 587–597. https://doi.org/10.1037/pst0000247.

Husby, V. M. 2025. "Challenge and Support: Scaffolding the Practicing Therapist in DP Supervision." *Journal of Clinical Psychology* 81, no. 5: 366–378. https://doi.org/10.1002/jclp.23774.

IBM Corp. 2023. *IBM SPSS Statistics for Windows* (Version 29.0.2.0) [Computer Software]. IBM Corp.

Janse, P. D., D. T. Bakker-Brehm, N. Geurtzen, A. Scholing, and G. J. Hutschemaekers. 2023. "Therapists' Self-Assessment of Time Spent on Learning Activities and Its Relationship to Treatment Outcomes: A Replication Study." *Journal of Clinical Psychology* 79, no. 4: 1070–1081. https://doi.org/10.1002/jclp.23459.

Levenson, H., V. Gay, and J. L. Binder. 2023. *Deliberate Practice in Psychodynamic Psychotherapy*. American Psychological Association. https://doi.org/10.1037/0000351-000.

Lipp, S. L. 2019. *The Effects of Case Conceptualization Training and Deliberate Practice Coaching on Counselor Competence* Dissertation. Florida Atlantic University. https://www.proquest.com/docview/2244327285/abstract/FA12FD9F1B6A49FBPQ/1.

MacCallum, R. C., K. F. Widaman, S. Zhang, and S. Hong. 1999. "Sample Size in Factor Analysis." *Psychological Methods* 4, no. 1: 84–99. https://doi.org/10.1037/1082-989X.4.1.84.

Mahon, D. 2023. "A Scoping Review of Deliberate Practice in the Acquisition of Therapeutic Skills and Practices." *Counselling and Psychotherapy Research* 23, no. 4: 965–981. https://doi.org/10.1002/capr.12601.

Mahon, D. 2024. "'Don't Just Magically Assume You're Going to Get Better as a Therapist as Time Goes on': An Exploration of the Use of Deliberate Practice." *Counselling and Psychotherapy Research* 24, no. 3: 913–924. https://doi.org/10.1002/capr.12746.

McClelland, D. C., J. W. Atkinson, and R. A. Clark. 2012. *The Achievement Motive*. Literary Licensing, LLC.

McLeod, J. 2022. "How Students Use Deliberate Practice During the First Stage of Counsellor Training." *Counselling and Psychotherapy Research* 22, no. 1: 207–218. https://doi.org/10.1002/capr.12397.

Miller, S. D., D. Chow, S. Malins, and M. A. Hubble, eds. 2023. *The Field Guide to Better Results: Evidence-Based Exercises to Improve Therapeutic Effectiveness*. 1st ed. American Psychological Association.

Miller, S. D., M. A. Hubble, and D. Chow. 2020. *Better Results: Using Deliberate Practice to Improve Therapeutic Effectiveness*. American Psychological Association. https://doi.org/10.1037/0000191-000.

Moosbrugger, H., and A. Kelava, eds. 2020. *Testtheorie und Fragebogenkonstruktion*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-61532-4.

Moulaert, V., M. G. M. Verwijnen, R. Rikers, and A. J. J. A. Scherpbier. 2004. "The Effects of Deliberate Practice in Undergraduate Medical Education." *Medical Education* 38, no. 10: 1044–1052. https://doi.org/10.1111/j.1365-2929.2004.01954.x.

Nurse, K., M. O'shea, M. Ling, N. Castle, and J. Sheen. 2024. "The Influence of Deliberate Practice on Skill Performance in Therapeutic Practice: A Systematic Review of Early Studies." *Psychotherapy Research* 35, no. 3: 353–367. https://doi.org/10.1080/10503307.2024.2308159.

Olsson, M., M. Gustafsson, J. Carlsson, T. Rousmaniere, S. Bergbom, and J. Norberg. 2025. "A Pilot Study of a Deliberate Practice Intervention for Therapist Trainees." *Nordic Psychology* 77, no. 2: 165–184. https://doi.org/10.1080/19012276.2024.2313532.

Ones, D. S., C. Viswesvaran, and A. D. Reiss. 1996. "Role of Social Desirability in Personality Testing for Personnel Selection: The Red Herring: Journal of Applied Psychology." *Journal of Applied Psychology* 81, no. 6: 660–679. https://doi.org/10.1037/0021-9010.81.6.660.

Orlowski, E. W., M. Friedlander, and M. Moeyaert. 2025. "Using Deliberate Practice for Inner Skills to Enhance Therapist In-Session Mindfulness and Presence." *Counselling Psychology Quarterly*: 1–20. https://doi.org/10.1080/09515070.2025.2551710.

Rammstedt, B., D. Danner, C. J. Soto, and O. P. John. 2020. "Validation of the Short and Extra-Short Forms of the Big Five Inventory-2 (BFI-2) and Their German Adaptations." *European Journal of Psychological Assessment* 36, no. 1: 149–161. https://doi.org/10.1027/1015-5759/a000481.

Rousmaniere, T. 2017. *Deliberate Practice for Psychotherapists: A Guide to Improving Clinical Effectiveness*. Routledge, Taylor & Francis Group.

Rousmaniere, T. 2024. *Deliberate Practice for Psychotherapists: A Guide to Improving Clinical Effectiveness*. 2nd ed. Routledge. https://doi.org/10.4324/9781003400622.

Schmidt, F. T. C., J. Fleckenstein, J. Retelsdorf, L. Eskreis-Winkler, and J. Möller. 2019. "Measuring Grit: A German Validation and a Domain-Specific Approach to Grit." *European Journal of Psychological Assessment* 35, no. 3: 436–447. https://doi.org/10.1027/1015-5759/a000407.

Schuler, H., and M. Prochaska. 2001. *Leistungsmotivationsinventar—Dimensionen Berufsbezogener Leistungsorientierung*. Hogrefe. https://www.hogrefe.com/at/shop/leistungsmotivationsinventar.html.

Tabachnick, B. G., and L. S. Fidell. 2007. *Using Multivariate Statistics*. 5th ed, Pearson internat. ed., [Nachdr.]. Pearson.

The Jamovi Project. 2024. "Jamovi (Version 2.3) [Computer Software]." https://www.jamovi.org/.

Tracey, T. J. G., B. E. Wampold, R. K. Goodyear, and J. W. Lichtenberg. 2015. "Improving Expertise in Psychotherapy." *Psychotherapy* 50, no. 1: 7–13.

Vaz, A., and T. Rousmaniere. 2022. *Clarifying Deliberate Practice for Mental Health Training*. Sentio University. http://www.sentio.org.

Wampold, B. E., and Z. E. Imel. 2015. *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work*. 2nd ed. Routledge/Taylor & Francis Group.

Wampold, B. E., and J. Owen. 2021. "Therapist Effects: History, Methods, Magnitude, and Characteristics of Effective Therapists." In *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change: 50th Anniversary Edition*, edited by M. Barkham, W. Lutz, and L. G. Castonguay, 7th ed., 297–326. John Wiley & Sons, Inc.

Westra, H. A., N. Norouzian, L. Poulin, et al. 2021. "Testing a Deliberate Practice Workshop for Developing Appropriate Responsivity to Resistance Markers." *Psychotherapy* 58, no. 2: 175–185. https://doi.org/10.1037/pst0000311.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** capr70181-sup-0001-AppendixS1.docx.