ORIGINAL ARTICLE

Comparing quality and engagement in face-to-face and online teacher professional development

Nina Mulaimović[1] | Eric Richter[1] | Rebecca Lazarides[1,2] | Dirk Richter[1]

[1]Department of Educational Science, University of Potsdam, Potsdam, Germany

[2]Cluster Science of Intelligence (SCIoI), Technische Universität Berlin, Berlin, Germany

Correspondence
Nina Mulaimović, Department of Educational Science, University of Potsdam, Karl-Liebknecht-Straße 24-25, Potsdam 14476, Germany.
Email: nina.mulaimovic@uni-potsdam.de

Abstract

In order for teachers to successfully gain new knowledge during professional development (PD), courses must be of high quality and stimulate active involvement from participants. More and more PD courses are taking place online, without clear evidence of whether face-to-face and online courses differ in terms of their quality or level of participants' engagement. The present study investigates differences between face-to-face and online PD with respect to certain quality characteristics: clarity and structure, cognitive activation, collaboration and practical relevance, as well as participants' behavioural, cognitive and affective engagement. The study is based on 2210 teachers from Germany who participated in 1 of 137 face-to-face or 54 online PD courses. Although participants rated face-to-face and online courses very positively regarding all quality characteristics and engagement dimensions, they evaluated online courses slightly less favourably compared to face-to-face courses. Implications for practice and research are derived to help ensure high-quality PD offerings in the future.

KEYWORDS
continuing education, effectiveness, participant engagement, PD quality characteristics, professionalization

Practitioner notes

What is already known about this topic

- Face-to-face and online PD have the potential to be similarly effective.
- PD quality and participants' engagement can be assumed to be predictors of PD effectiveness.
- PD quality contains clarity and structure, cognitive activation, collaboration and practical relevance.
- Engagement is a three-dimensional construct composed of behavioural, cognitive and affective components.

What this paper adds

- PD quality was rated very positively for online and face-to-face courses.
- Participants rated the quality of online PD lower compared to face-to-face PD.
- Participants rated their engagement in online PD lower compared to face-to-face PD.

Implications for practice and/or policy

- PD format should always be chosen with which a higher benefit can be achieved.
- Quality assurance should take place before PD is conducted.

INTRODUCTION

Effective teacher PD can be defined 'as structured professional learning that results in changes to teacher knowledge and practices, and improvements in student learning outcomes' (Darling-Hammond et al., 2017, p. 2). It can be provided in at least two ways: face-to-face or online. In face-to-face formats, teachers interact with the PD material, colleagues and the facilitator in person. In contrast, online formats allow interaction between individuals through the Internet (Fishman, 2016). This can include synchronous or asynchronous activities or combinations of both. Synchronous online learning activities take place in real time and are similar to learning activities in face-to-face settings, in that the facilitator supervises one or more teachers who are simultaneously participating in a virtual learning environment (eg, distance education courses offered by universities; Bates et al., 2016). Meanwhile, asynchronous online learning activities occur at different times for different participants, allowing them to determine what, how and when they learn (eg, teacher social networks; Bates et al., 2016). Since face-to-face and synchronous online PD show the greatest similarity, the comparison in the present study refers to these two PD formats.

One question that currently remains unanswered in PD research has to do with the factors that contribute to the effectiveness of PD activities. In this context, PD quality has emerged as one possible predictor of teachers' learning and students' achievement (eg, Desimone & Garet, 2015; Fischer et al., 2018). Although there is no uniform procedure for measuring the quality of PD courses so far (Desimone, 2009), researchers have already been able to identify some quality characteristics that have a positive impact on teachers' knowledge and practices, as well as their students' learning outcomes (eg, Meyer et al., 2023; Darling-Hammond et al., 2017; Desimone, 2009). Some studies refer to face-to-face PD while others refer to online PD, but the investigated quality characteristics seem to be very similar. However, it is still unclear whether these characteristics are expressed differently in the two PD formats.

Teacher engagement, on the other hand, can be assumed to be another predictor of the effectiveness of PD activities, as it 'is critical for learning processes and outcomes and affects teachers' professional learning, and, hence, their teaching' (Vivante & Vedder-Weiss, 2023, p. 1). Researchers largely agree on how to capture engagement in face-to-face and online courses (eg, Martin & Borup, 2022; Vivante & Vedder-Weiss, 2023), even though most PD research to date has examined engagement in face-to-face formats (eg, Jansen in de Wal et al., 2014; Ryan, 2019). Accordingly, it remains unanswered to date whether teacher engagement differs in face-to-face and online PD.

The present study addresses the research desiderata described above and aims to expand the basic understanding of online PD compared to face-to-face PD. To this end, the study focuses on PD quality and teacher engagement in PD activities, two aspects that are considered central to the effectiveness of teacher PD.

LITERATURE REVIEW

Quality characteristics of effective teacher professional development

High-quality PD can have a positive impact on the effectiveness of the course (eg, Desimone & Garet, 2015; Fischer et al., 2018). So far, research has not settled on a uniform set of quality characteristics of effective teacher PD. However, when summarizing the most frequently mentioned aspects, it can be stated that effective teacher PD should be well structured, include activating tasks, provide opportunities for exchange and make connections to one's professional practice. For this reason, we identify the characteristics of clarity and structure, cognitive activation, collaboration and practical relevance as key features of high-quality teacher PD.

Clarity and structure refer to precisely defined learning goals, an understandable explanation of the PD content, an efficient use of time and a transparent and reasonable course of the PD (cf., Cordingley et al., 2015; Lipowsky & Rzejak, 2021; Sims et al., 2021). For example, facilitators should set clear expectations and learning goals when designing PD (van Nuland et al., 2020) and choose appropriate pedagogical approaches (ie, meaningful structure and course components) to help teachers achieve them (Carrillo & Flores, 2023). *Cognitive activation* involves incorporating prior knowledge, providing thought-provoking questions and tasks and providing feedback and opportunities to reflect on one's professional activities (cf., Borko et al., 2010; Darling-Hammond et al., 2017; Garet et al., 2001). In this context, the term 'active learning' is often used, a teaching approach that should allow teachers to try out and reflect on newly acquired strategies. Furthermore, *collaboration* includes the promotion of information exchange (cf., Butler et al., 2004; Didion et al., 2020; Hubers et al., 2020). For example, joint group activities can give teachers the opportunity to develop ideas for their professional practices, share experiences and clarify misunderstandings (Carrillo & Flores, 2023; Lefstein et al., 2020; Onrubia et al., 2022). Finally, *practical relevance* contains all aspects related to the relationship between the PD content and one's professional practice, such as aligning the PD content with discipline-specific curricula and using practice-based models (eg, in the form of teaching or student work samples; cf., Darling-Hammond et al., 2017; Desimone, 2009; Lipowsky & Rzejak, 2021).

The characteristics described above predominantly originate from studies that investigated face-to-face PD. However, there is already some research showing that these characteristics also apply to online PD. For example, Meyer et al. (2023) identified clarity and structure, cognitive activation and collaboration as predictors of teachers' satisfaction with online PD, while also showing that cognitive activation and collaboration predicted changes in their professional practice. The finding that collaboration is an important feature of

high-quality online PD also emerges from a systematic review by Bragg et al. (2021). They note that teachers' interactions during online PD activities can increase their satisfaction with online PD and may be associated with changes in their beliefs about teaching as well as their instructional practices. Furthermore, other research suggests a relationship between the perceived clarity of online PD structures and teachers' satisfaction with online PD (eg, Reeves & Pedulla, 2011).

Although these studies show that the identified quality characteristics also have a positive impact on teachers' knowledge and instructional practices in the context of online PD, it is likely that designing online courses poses particular challenges for facilitators. For example, we already know from higher education research that online learning environments increase the likelihood of learners' minds wandering (Hollis & Was, 2016), so facilitators may feel particularly challenged to maintain the participants' cognitive activation. Moreover, it may be more difficult for facilitators to ensure collaboration among participants because online courses do not allow for physical proximity (Al-Naabi, 2023). This problem has already been confirmed by participants as a particularly great challenge of online PD (Collins & Liang, 2015; Howard, 2021; Wynants & Dennis, 2018). For example, online PD is limited in its ability to encourage and foster professional interaction, learning communities and information sharing (Collins & Liang, 2015).

In order to investigate whether face-to-face and online PD have comparable levels of effectiveness, notwithstanding the particular challenges of online PD, courses with similar content but different delivery formats (face-to-face and online) were developed and compared with respect to their learning outcomes. No significant differences emerged in terms of teachers' positive perceptions of the PD (Yoon et al., 2020), their development of knowledge, beliefs and instructional practices (Fishman et al., 2013; Russell et al., 2009), or their students' learning outcomes (Fishman et al., 2013; Yoon et al., 2020). Consequently, it can be stated that face-to-face and online PD have the potential to be similarly effective. This finding is also consistent with equivalence theory (Simonson, 1999; Simonson et al., 1999), which assumes that face-to-face and online learners achieve equivalent learning outcomes when equivalent learning experiences are provided. However, it is questionable whether this finding can also be shown in non-controlled settings, that is, whether it can also be transferred to the broad spectrum of PD courses. So far, mainly PD outcomes rather than PD quality have been compared, the latter of which we would like to address with the present study.

Engagement in teacher professional development

The effectiveness of PD can be predicted not only by the quality of the course but also with a high degree of probability by the engagement of the participants (Vivante & Vedder-Weiss, 2023), as we already know from studies on adult learning that learners' engagement is critical for successful learning processes and outcomes (McDonough, 2014; Ornelles et al., 2019).

Engagement is generally understood as a multidimensional construct composed of behavioural, cognitive and affective (also known as emotional) components (Fredricks et al., 2004). *Behavioural engagement* describes involvement in tasks (eg, asking questions, completing work assignments, showing effort and persistence; Fredricks et al., 2004). *Cognitive engagement* concerns the mental effort and investment in intellectual challenges (eg, developing strategies and solving problems), and *affective engagement* includes positive and negative feelings towards the instructor, peers, work assignments or the institution (eg, joy, excitement, anxiety and boredom; Fredricks et al., 2004).

The dimensions of engagement can be investigated in both face-to-face and online courses, although they may be less pronounced online (Martin & Borup, 2022). For example, *cognitive engagement* in online courses might be limited by the fact that first-time online learners not only need to learn the subject-matter content but also how to learn online (Lowes & Lin, 2015). In addition, online courses could have a negative impact on *affective engagement*, as participation in videoconferencing may be associated with a feeling of mental exhaustion after some time (Bailenson, 2021). This hypothesis is already supported by findings from higher education research showing that students' affective engagement in online courses is significantly lower than in face-to-face courses (Salta et al., 2022). However, the decrease in affective engagement in this study can largely be explained by the simultaneous decrease in the level of human interaction (either student–student or student-instructor) during the transition from face-to-face to online courses. Finally, with regard to *behavioural engagement* in online courses, it could be critical that learners tend to engage in off-task behaviours when learning activities do not require their active involvement (Martin & Borup, 2022). For example, learners participating in massive open online courses (MOOCs) tend to stop watching learning videos after only a few minutes (Guo et al., 2014). However, learners might also be prone to distraction in videoconferences when little attention is required from them (Martin & Borup, 2022).

So far, there have only been a few studies on the engagement of teachers in PD, and the studies that do investigate engagement mainly refer to face-to-face PD. For example, Jansen in de Wal et al. (2014) investigated the extent to which teachers' affiliation with one of four profiles of the motivational dimensions from self-determination theory ('extremely autonomous', 'moderately motivated', 'highly autonomous' and 'externally regulated') was related to their engagement in face-to-face PD. They showed that teachers with profiles of identified regulation and intrinsic motivation were more likely to engage in PD activities. In contrast, Ryan (2019) conducted a naturalistic observation study to analyse teachers' behaviour in face-to-face PD. He showed that participants were mostly distracted, for example, by communicating with others, looking around the room or playing with objects, rather than attentively and actively involved in the course. Finally, Vivante and Vedder-Weiss (2023) extended the understanding of engagement by developing a multimodal and situated analysis tool, which captures engagement by the expected behaviour based on both body language and verbal expressions. An initial application of the analysis tool in the context of different face-to-face PD activities showed that teachers were engaged most often 'as expected' verbally and non-verbally during game playing and inquiry into practice. In contrast, during learning in the whole group, although they were verbally engaged most often 'as expected,' their non-verbal engagement was both below and above expectations with about equal frequency.

THE PRESENT STUDY

The effectiveness of teacher PD is significantly related to the quality of the offering (eg, Desimone & Garet, 2015; Fischer et al., 2018) and most likely also to the engagement of the participants (Vivante & Vedder-Weiss, 2023). Given the two primary modalities for delivering impactful teacher PD—face-to-face versus online—each characterized by distinct modes of interaction (physical presence vs. virtual connectivity), it becomes compelling to investigate potential disparities across these dimensions. Experimental comparisons between face-to-face and online PD have already shown that both formats can be similarly effective (Fishman et al., 2013; Russell et al., 2009; Yoon et al., 2020). However, it is questionable whether this is also the case in non-controlled settings, that is, whether this finding can also be transferred to the broad spectrum of PD courses. Furthermore, so far, mainly PD outcomes rather

than PD quality and participants' engagement have been compared, which we would like to change with the present study. This results in the following two research questions:

Research question 1: How does the quality of teacher professional development differ in terms of clarity and structure, cognitive activation, collaboration and practical relevance between face-to-face and synchronous online courses?

Considering previous research (eg, Al-Naabi, 2023; Hollis & Was, 2016), we assume that teachers rate the quality of online PD (esp. with regard to cognitive activation and collaboration) lower than that of face-to-face PD.

Research question 2: How does teachers' behavioural, cognitive and affective engagement differ between face-to-face and synchronous online professional development?

Taking into account previous research (eg, Martin & Borup, 2022; Salta et al., 2022), we assume that teachers rate their engagement (behavioural as well as cognitive and affective) lower in online PD than in face-to-face PD.

METHOD

Study design and sample

The present study is a cross-sectional quantitative questionnaire study that was conducted online in a large federal state in Germany between October 2021 and July 2022. In the course of this, in-service teachers assessed both the quality and their level of engagement in 1 of 191 different regular PD courses offered by the state PD provider. This provider also supported the evaluation process by allowing time for participation in the survey at the end of the courses.

Of the 191 PD courses, 137 were face-to-face and 54 were synchronous online courses. All courses vary greatly in their content and target groups, as both formats (face-to-face and online) address different topic areas (eg, subject didactics, school development and classroom development; see Appendix, Figure A1) and teachers from different types of schools and grades. As a result, we were able to reflect the diversity of the PD landscape in our study in the best possible way.

At the end of each PD course, participants were invited by the facilitator to evaluate the course as part of a written online survey. Therefore, every teacher in the PD had the opportunity to participate voluntarily in the study during the regular course time. A total of 2210 teachers completed the questionnaire. However, at this point, we cannot provide information about the response rate, as we do not have any information about how many teachers took part in the courses for data protection reasons. Yet, we can point out that on average, 12 individuals participated in the study in each PD course. This is an acceptable number, as it corresponds to findings of programme analysis for PD courses in one German federal state, in which approximately 10–12 teachers participated per course (Richter et al., 2020).

The majority of the surveyed teachers identified themselves as female (72.1%) and had completed a traditional teacher training programme (82.2%). In Germany, this comprises a bachelor's and master's degree and a mandatory induction phase in a school completed with an exam. In addition, the majority of teachers in this sample had 5 or more years of teaching experience (77.5%) and worked full-time (67.6%). About 30.5% of the teachers were working in primary schools, all others in secondary schools.

To examine whether teachers who participate in face-to-face and online courses differ with regard to sociodemographic and occupational characteristics, a multivariate ANOVA was conducted (Table 1). We found that teachers participating in online courses were significantly more likely to be female ($F = 4.94$; $p = 0.03$; $\eta^2 = 0.05$), significantly more likely to have completed a traditional teacher training programme ($F = 11.76$; $p < 0.01$; $\eta^2 = 0.10$) and

TABLE 1 Descriptive statistics of sociodemographic and occupational characteristics of teachers participating in face-to-face and online PD.

Variable	Face-to-face PD		Online PD		Group comparison		
	M	*SD*	*M*	*SD*	*F*	*p*	η^2
Gender (proportion female)[a]	74.76	24.16	84.75	17.60	4.94	0.03	0.05
Scope of employment (proportion part-time)[b]	34.85	24.26	33.22	23.53	0.11	0.74	<0.01
Teaching experience (proportion 5 or more years)[a]	68.17	26.58	80.62	16.44	6.75	0.01	0.06
Teacher certification (proportion traditional)[b]	77.55	22.69	91.31	12.34	11.76	<0.01	0.10
School type (proportion primary school)[a]	46.48	45.63	38.94	40.14	0.72	0.40	<0.01

Note: The reported values represent aggregated values at the group level. Means and standard deviations are given as percentages.

[a]Face-to-face: $n=137$, online: $n=54$.

[b]Face-to-face: $n=71$, online: $n=37$ (variables marked with b were not collected in all courses).

significantly more likely to have 5 or more years of teaching experience ($F=6.75$; $p=0.01$; $\eta^2=0.06$). In contrast, no significant differences were found with respect to the scope of employment or the type of school. Accordingly, teachers' gender, teacher certification and teaching experience were included as covariates in the analyses.

Measures

Professional development quality

The teachers evaluated the PD quality with a new instrument (see Table A1), which was developed in a multi-stage process by stakeholders from educational research, practice and administration (Richter & Richter, 2024). The first step was to design the instrument. To this end, a joint workshop was held in which key quality characteristics of effective teacher PD were identified on the basis of previous research findings. It was possible to agree on the four quality dimensions (clarity and structure, cognitive activation, collaboration and practical relevance) already presented in the literature review. The identified quality characteristics were then operationalized with items by a smaller multidisciplinary group before feedback was obtained and directly incorporated in a further workshop with the larger group. In the second step, the instrument was piloted. The piloting took place from October 2021 to July 2022, with two waves of data collection in which the instrument was shared with different teachers.

The instrument has a high psychometric quality and the associated model has an acceptable model fit: $\chi^2(274)=785.01$, CFI$=0.95$, RMSEA$=0.02$, SRMR between$=0.12$, SRMR within$=0.04$ (Richter & Richter, 2024). All factor loadings are within an acceptable range, both for the level 1 variables (participant level; 0.63–0.73 for clarity and structure, 0.61–0.71 for cognitive activation, 0.67–0.83 for collaboration and 0.70–0.76 for practical relevance) and the level 2 variables (course level; 0.86–0.96 for clarity and structure, 0.74–0.96 for cognitive activation, 0.72–0.99 for collaboration and 0.90–0.99 for practical relevance). Clarity and structure of the course were rated using five items (eg, 'The goals of the event were clearly stated'). We calculated the internal consistency of the scale using McDonald's ω (McDonald, 1999), since this value, unlike Cronbach's α, takes into account varying item loadings and varying error variances (cf., Goodboy & Martin, 2020; Hayes & Coutts, 2020). McDonald's ω can be interpreted like Cronbach's α, that is, acceptable internal consistency

is associated with a value of 0.70 or higher (McNeish, 2018). For the scale of clarity and structure, a $\omega = 0.84$ was obtained. Moreover, we determined the number of missing values per scale. The scale of clarity and structure had no missing values. Cognitive activation in the PD was assessed using six items (eg, 'In the event, questions or tasks were asked that made me think'; $\omega = 0.86$; 0.6% missing values), collaboration using three items (eg, 'During the event, I had the opportunity to discuss the event content in depth with the other participants'; $\omega = 0.83$; 0.8% missing values) and practical relevance using four items (eg, 'The aspects covered in the event were related to my current professional practice'; $\omega = 0.86$; 1.0% missing values). All items were rated on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree).

Engagement in professional development

The instrument for assessing teachers' engagement in the PD courses is based on an instrument from Chan et al. (2023), as it is composed of previously established, repeatedly used and reliable constructs. For the purposes of this study, we have translated it into German and modified it according to the study context. According to this instrument, teachers rated their engagement along three dimensions: behavioural, cognitive and affective. Here, the assessment of behavioural engagement was based on five items (eg, 'I have asked questions'). One item was excluded from the analyses due to insufficient reliability, so ultimately four items were used and a $\omega = 0.81$ was obtained (4.5% missing values). The cognitive engagement was assessed using four items (eg, 'During the event, I thought deeply about the content'). Here, two items were removed due to insufficient reliability, so ultimately two items were used and a $\omega = 0.65$ was obtained (4.7% missing values). Finally, teachers assessed their affective engagement based on four items (eg, 'I am glad I participated in the event'; $\omega = 0.85$; 5.1% missing values). All items were rated on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree).

Statistical analysis

Multilevel analyses were conducted due to the clustered data structure that assigns teachers to PD groups. Thereby, doubly latent models (latent variable and latent aggregation) were used as proposed by Marsh et al. (2009) because they take into account measurement error at levels 1 and 2 (based on multiple level 1 indicators) and level 2 sampling error due to the aggregation from levels 1 and 2. The assumptions for this according to Morin et al. (2014) are met. All models were computed using maximum-likelihood estimation (MLE; Myung, 2003) in Mplus 8.3 (Muthén & Muthén, 1998–2017). We evaluated model fit using the recommendations of Hu and Bentler (1999) for comparative fit index (CFI > 0.95), Tucker–Lewis index (TLI > 0.95), root-mean-squared error of approximation (RMSEA < 0.06) and standardized root-mean-squared residual (SRMR < 0.08).

To answer the first research question, a regression model was estimated that simultaneously considered all characteristics of PD quality (clarity and structure, cognitive activation, collaboration and practical relevance) as dependent variables (model 1). Additional analyses showed that the parameter estimates of this model did not differ substantially from models that included each variable separately. Moreover, teachers' gender, teacher certification and teaching experience were included as covariates at the teacher level (level 1), and aggregates of these variables and PD format were included as covariates at the group level (level 2).

To answer the second research question, the regression model was modified in that all dimensions of engagement in the PD (behavioural engagement, cognitive engagement and affective engagement) were included simultaneously as dependent variables instead of the PD quality characteristics (model 2). Again, the overall model was not significantly different from models with single variables.

The data set has missing values due to non-existent item responses. To account for these nonetheless, the full-information maximum-likelihood (FIML; Enders, 2001) procedure was used, which is preferred to alternative approaches such as listwise or pairwise deletion (Graham, 2012).

RESULTS

Quality of face-to-face and online teacher professional development

Before we present the results of the doubly latent multilevel analyses, which allow an appropriate comparison of quality between face-to-face and online PD (research question 1), we describe the descriptive statistics of the manifest scores (Table 2). The statistics allow us to evaluate the overall agreement with the items used in the survey instrument. We find that all quality characteristics obtained very positive ratings that greatly exceed the theoretical mean of the scale. This means that participants tend to report high clarity and structure, cognitive activation, collaboration and practical relevance overall.

To compare the means of the quality characteristics, we estimated a structural equation model that accounts for the nested structure of the data and uses latent variables to describe the quality characteristics. The results of the multilevel analyses are presented in Table 3 and include y-standardized regression coefficients in order to compare results between the models. We find that the two PD formats differ significantly with regard to clarity and structure ($\hat{\beta} = -0.31$, $p < 0.01$), cognitive activation ($\hat{\beta} = -0.38$, $p < 0.01$), collaboration ($\hat{\beta} = -0.51$, $p < 0.01$) and practical relevance ($\hat{\beta} = -0.24$, $p < 0.05$). Thus, compared to face-to-face courses, teachers rate online courses as less transparent and structured, less cognitively activating, less collaborative and less practically relevant. According to Cohen (1988), the effect of collaboration can be classified as medium, whereas the effects of clarity and structure, cognitive activation and practical relevance are small. In addition, it appears that teachers' teaching experience is significantly positively related to clarity and structure ($\hat{\beta} = 0.10$, $p < 0.01$) and cognitive activation ($\hat{\beta} = 0.06$, $p < 0.05$). That is, teachers with 5 or more years of teaching experience generally perceive PD courses as more transparent and structured, as well as more cognitively activating, than teachers with <5 years of teaching

TABLE 2 Descriptive statistics of quality characteristics in face-to-face and online PD averaged for all PD courses.

Variable	Face-to-face PD[a]		Online PD[b]	
	M	*SD*	*M*	*SD*
Clarity and structure	3.64	0.28	3.47	0.41
Cognitive activation	3.39	0.32	3.19	0.41
Collaboration	3.60	0.37	3.18	0.58
Practical relevance	3.55	0.30	3.44	0.42

Note: The reported values represent aggregated values at the group level.

[a]Face-to-face: $n = 137$.

[b]Online: $n = 54$.

TABLE 3 Results of the multilevel regression analyses: PD format and teacher characteristics as predictors of PD quality (*y*-standardized regression coefficients).

Predictors	Clarity and structure $\widehat{\beta}$ (*SE*)	Cognitive activation $\widehat{\beta}$ (*SE*)	Collaboration $\widehat{\beta}$ (*SE*)	Practical relevance $\widehat{\beta}$ (*SE*)
Group level				
PD format (1 = online)	−0.31** (0.10)	−0.38** (0.08)	−0.51** (0.08)	−0.24* (0.09)
Gender (1 = female)	0.13 (0.09)	0.11 (0.08)	−0.02 (0.08)	0.10 (0.09)
Teacher certification (1 = traditional)	−0.02 (0.14)	0.05 (0.18)	0.03 (0.17)	0.08 (0.18)
Teaching experience (1 = 5 or more years)	−0.10 (0.15)	0.11 (0.12)	0.08 (0.09)	−0.18 (0.11)
Teacher level				
Gender (1 = female)	0.03 (0.03)	0.02 (0.03)	−0.01 (0.03)	0.05 (0.03)
Teacher certification (1 = traditional)	<−0.01 (0.04)	0.02 (0.04)	0.05 (0.05)	0.03 (0.04)
Teaching experience (1 = 5 or more years)	0.10** (0.03)	0.06* (0.03)	0.05 (0.03)	0.04 (0.03)
R^2 (between)	0.13	0.17	0.27	0.11
R^2 (within)	0.01	0.01	<0.01	0.01

Note: The reported values for gender, teacher certification and teaching experience at the group level represent aggregated values. $\chi^2 = 1106.05$, $df = 379$, $p = <0.01$, CFI = 0.94, TLI = 0.93, RMSEA = 0.03, SRMR between = 0.12 and SRMR within = 0.04. The dependent variables show substantial variance between the level 2 units: clarity and structure: ICC = 0.28, cognitive activation: ICC = 0.24, collaboration: ICC = 0.37 and practical relevance: ICC = 0.25.

*$p < 0.05$; **$p < 0.01$.

experience. However, only very small effects are found here. A regression model without control variables can be found in the appendix (see Table A2). Here, the main findings related to the quality differences do not substantially differ.

Engagement in face-to-face and online teacher professional development

Before addressing research question 2, we describe the descriptive statistics of the manifest means for all three engagement dimensions (Table 4). We find that all observable scores are above the theoretical scale mean and all scores are above the 3.0 value. This means that participants tend to report high behavioural, cognitive and affective engagement overall. A direct comparison of these values between the individual engagement scales is not possible because different items were used.

In the next step, we examine the differences between the PD formats on the basis of a doubly latent multilevel structural equation model. The results of this model show that participants' engagement differs as a function of the PD format, in the sense that participants in online PD show lower engagement (Table 5). For behavioural ($\widehat{\beta} = -0.45$, $p < 0.01$) and cognitive engagement ($\widehat{\beta} = -0.52$, $p < 0.01$), we find medium effect sizes according to Cohen (1988), whereas for affective engagement, we find small effects ($\widehat{\beta} = -0.31$, $p < 0.01$). In addition, it appears that teachers' certification ($\widehat{\beta} = 0.10$, $p < 0.05$) and teaching experience ($\widehat{\beta} = 0.08$, $p < 0.05$) are significantly positively related to behavioural engagement. That is, teachers with traditional teacher certification and 5 or more years of teaching experience

TABLE 4 Descriptive statistics of engagement dimensions in face-to-face and online PD averaged for all PD courses.

Variable	Face-to-face PD[a]		Online PD[b]	
	M	*SD*	*M*	*SD*
Behavioural engagement	3.47	0.29	3.23	0.38
Cognitive engagement	3.47	0.27	3.27	0.35
Affective engagement	3.46	0.34	3.27	0.43

Note: The reported values represent aggregated values at the group level.

[a]Face-to-face: $n = 137$.

[b]Online: $n = 54$.

TABLE 5 Results of the multilevel regression analyses: PD format and teacher characteristics as predictors of PD engagement (*y*-standardized regression coefficients).

Predictors	Behavioural engagement	Cognitive engagement	Affective engagement
	$\hat{\beta}$ *(SE)*	$\hat{\beta}$ *(SE)*	$\hat{\beta}$ *(SE)*
Group level			
PD format (1 = online)	−0.45** (0.09)	−0.52** (0.08)	−0.31** (0.09)
Gender (1 = female)	−0.18* (0.09)	0.02 (0.10)	0.02 (0.09)
Teacher certification (1 = traditional)	0.24 (0.12)	0.27 (0.21)	0.19 (0.17)
Teaching experience (1 = 5 or more years)	−0.02 (0.10)	−0.05 (0.17)	−0.05 (0.12)
Teacher level			
Gender (1 = female)	0.01 (0.03)	0.09* (0.03)	0.09** (0.03)
Teacher certification (1 = traditional)	0.10* (0.04)	−0.02 (0.05)	−0.08 (0.04)
Teaching experience (1 = 5 or more years)	0.08* (0.03)	0.07 (0.04)	0.03 (0.03)
R^2 (between)	0.29	0.34	0.14
R^2 (within)	0.02	0.01	0.02

Note: The reported values for gender, teacher certification and teaching experience at the group level represent aggregated values. $\chi^2 = 655.69$, $df = 129$, $p = <0.01$, CFI = 0.92, TLI = 0.90, RMSEA = 0.04, SRMR between = 0.13 and SRMR within = 0.05. The dependent variables show substantial variance between the level 2 units: behavioural engagement: ICC = 0.18, cognitive engagement: ICC = 0.15 and affective engagement: ICC = 0.29.

*$p < 0.05$; **$p < 0.01$.

generally rate their behavioural engagement in PD courses higher than teachers with non-traditional teacher certification and <5 years of teaching experience. However, only very small effects are found here. Furthermore, some significant effects emerge with respect to the gender of the teachers. For example, at the teacher level (level 1), gender is significantly positively related to cognitive ($\hat{\beta} = 0.09$, $p < 0.05$) and affective engagement ($\hat{\beta} = 0.09$, $p < 0.01$), whereas at the group level (level 2), it is significantly negatively related to behavioural engagement ($\hat{\beta} = -0.18$, $p < 0.05$). This means, on the one hand, that female teachers generally rate their cognitive and affective engagement in PD courses higher than male teachers and, on the other hand, that female teachers in the group rate their

behavioural engagement lower compared to their male colleagues. However, again the effects are very small to small. A regression model without control variables can be found in the appendix (see Table A3). Here, the main findings related to the engagement differences do not substantially differ.

DISCUSSION

Teacher PD provides valuable learning opportunities that can have a positive impact on teachers' professional practice and, in turn, on students' learning success (eg, Desimone & Garet, 2015; Fischer et al., 2018). To achieve this, PD activities should be of high quality and show a high level of participants' engagement (eg, Darling-Hammond et al., 2017; Vivante & Vedder-Weiss, 2023). However, it has not yet been systematically investigated whether face-to-face and online PD differ in these characteristics, which was the aim of the present study.

We were able to show that teachers rate the quality of and their engagement in the courses very positively in both face-to-face and online PD. Online PD can therefore certainly keep up and appears to be an attractive alternative to face-to-face PD. However, we also find that teachers rate both the quality of and their engagement in online PD significantly lower than in face-to-face PD. This finding is not in line with previous research findings showing that both PD formats have the potential to be similarly effective (Fishman et al., 2013; Russell et al., 2009; Yoon et al., 2020). Here, however, it is important to note that previous research has predominantly taken place in controlled settings where a standardized PD programme was developed and implemented in both PD formats, whereas our study has higher ecological validity: It took place in a naturalistic setting and compared a wide range of PD courses with different content and facilitators. In addition, previous studies have focused on PD outcomes rather than PD quality and participants' engagement. It is, therefore, quite likely that online PD courses perform worse than face-to-face PD courses.

With regard to PD *quality*, we showed that teachers rate online PD as less transparent and structured, less cognitively activating, less collaborative and less practically relevant in contrast to face-to-face PD. This showed a medium effect for collaboration and small effects for clarity and structure, cognitive activation and practical relevance. At this point, it is not possible to provide explanations for this phenomenon. We suspect, however, that these differences are largely due to the novelty of online PD so it remains questionable how well-prepared facilitators are and how particular challenges should be addressed. In this regard, there is already some research evidence suggesting that facilitators often struggle to model effective integration of technology, and consequently, need pedagogical support in creating digital teaching courses (Amhag et al., 2019; Becuwe et al., 2017). In particular, it may be challenging for facilitators to maintain participants' cognitive activation because, as we already know from higher education research, online learning environments increase the likelihood of learners' minds wandering (Hollis & Was, 2016). Furthermore, digital participation goes along with physical distance (Al-Naabi, 2023), which means that facilitators may find it more difficult to create sufficient opportunities for interaction. The observation that the latter is perceived as one of the biggest challenges of online PD (Collins & Liang, 2015; Howard, 2021; Wynants & Dennis, 2018) may also be reflected in our results, which show the largest effect for collaboration. Moreover, the differences in PD quality may also be attributable to the different PD content, which we did not consider in more detail in our analyses.

With regard to PD *engagement*, we found that teachers rate their behavioural, cognitive and affective engagement in online courses lower than in face-to-face courses, showing medium effects for behavioural and cognitive engagement and a small effect for affective engagement. Again, we cannot provide explanations for this result based on our data, but we assume that the differences can largely be attributed to the particular challenges of online PD.

For example, cognitive engagement may be lower in online PD because someone who is new to learning online not only has to deal with the content but also with handling the technology used (Lowes & Lin, 2015). This could even lead to cognitive overload if, for example, people are asked to participate in a survey as part of an online session but have difficulty operating the digital survey tool. Furthermore, affective engagement might be lower because video-conferencing makes one feel mentally exhausted after some time (Bailenson, 2021), and behavioural engagement might be lower because one tends to engage in off-task behaviour when videoconferencing does not require active involvement (Martin & Borup, 2022). The engagement differences found may also be attributable to the limited face-to-face contact in online courses, which Salta et al. (2022) were already able to confirm for students' affective engagement, or to the different PD content, as we have already mentioned with respect to the quality differences. Moreover, we already know that teachers have different motivations for attending face-to-face and online PD courses, which could also provide a possible explanation for the differences between the formats (Fütterer et al., 2024).

Although we were able to show that teachers' attendance at PD courses varied by their gender, their teacher certification and their teaching experience, there are no to weak effects of these variables on the perceived quality and the self-assessed engagement in the courses. This finding is consistent with previous research, which, for example, has also found only a small gender-specific effect with regard to students' engagement in face-to-face and online courses (Salta et al., 2022).

The present study has not only several strengths but also some limitations that need to be discussed. For example, it was one of the first studies to investigate in an ecologically valid setting whether face-to-face and online PD courses differ in terms of their quality or level of participants' engagement. A large number of PD courses with different content and facilitators were considered, and opinions were gathered from a large and random sample of teachers. However, we have to acknowledge that our results are based on a German sample so replications are needed in other countries. Second, we relied exclusively on participants' self-reports and did not consider the PD content or the exact context in our analyses. Third, we did not measure baseline levels of teachers' engagement prior to the PD activities. Consequently, we cannot make any statement about whether the very engaged teachers are not simply found more frequently in face-to-face PD courses because they are more willing to overcome spatial barriers, or whether there really is an increase in engagement. Finally, the way in which the study was conducted could have led to a systematic bias in the group of study participants. Since participation was only possible once, at the end of a particular PD course, it is conceivable that people who were already less engaged in the course may also have chosen not to participate in the study. However, as we cannot determine the response rate, we cannot estimate the extent of the bias.

Notwithstanding the reported limitations, some implications for practice and research can be derived from our study. For example, despite the slightly better performance of face-to-face PD, the fact that both PD formats are rated very highly should not be disregarded when choosing the PD format. Additionally, it should not be forgotten that online PD has some advantages over face-to-face PD. As it is location independent, it can reach geographically isolated teachers or teachers with limited time resources (Bates et al., 2016; Dede et al., 2009; Fishman et al., 2013). Furthermore, it is a more sustainable and cost-effective alternative to face-to-face PD (Fishman et al., 2013). Accordingly, it should always be carefully considered with which format a higher benefit can be achieved. In any case, we recommend continuously evaluating the quality of the PD courses and ensuring that the quality characteristics of clarity and structure, cognitive activation, collaboration and practical relevance are met. Furthermore, it may be useful to prepare facilitators to deal with particular challenges of online PD and to familiarize them with suitable measures to counteract these. In this context, it would also be interesting to investigate whether facilitators need specific competencies to

conduct high-quality online PD and which facilitators offer high-quality online PD. Moreover, it seems particularly promising to investigate the extent to which the PD quality and the engagement of the participants are connected and whether losses in these two characteristics are also associated with losses in learning success.

Overall, the present study provides new insights for PD practice and research that can help ensure high-quality PD offerings in both face-to-face and online formats in the future.

CONFLICT OF INTEREST STATEMENT
We declare that there are no known conflicts of interest associated with this research.

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

ETHICS STATEMENT
This study adhered to all national and international regulations for protecting human subjects.

ORCID
Nina Mulaimović https://orcid.org/0000-0002-0367-4487
Eric Richter https://orcid.org/0000-0003-3572-2022
Rebecca Lazarides https://orcid.org/0000-0003-0392-4981
Dirk Richter https://orcid.org/0000-0002-2384-1588

REFERENCES
Al-Naabi, I. S. (2023). A phenomenographic exploration of the factors influencing teacher social presence in professional development webinars. *Journal of Education and Learning*, *12*(6), 18–27. https://doi.org/10.5539/jel.v12n6p18

Amhag, L., Hellström, L., & Stigmar, M. (2019). Teacher educators' use of digital tools and needs for digital competence in higher education. *Journal of Digital Learning in Teacher Education*, *35*(4), 203–220. https://doi.org/10.1080/21532974.2019.1646169

Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior*, *2*(1). https://doi.org/10.1037/tmb0000030

Bates, M. S., Phalen, L., & Moran, C. (2016). Online professional development: A primer. *Phi Delta Kappan*, *97*(5), 70–73. https://doi.org/10.1177/0031721716629662

Becuwe, H., Pareja Roblin, N., Tondeur, J., Thys, J., Castelein, E., & Voogt, J. (2017). Conditions for the successful implementation of teacher educator design teams for ICT integration: A Delphi study. *Australasian Journal of Educational Technology*, *33*(2), 159–172. https://doi.org/10.14742/ajet.2789

Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. *International Encyclopedia of Education*, *7*(2), 548–556. https://doi.org/10.1016/B978-0-08-044894-7.00654-0

Bragg, L. A., Walsh, C., & Heyeres, M. (2021). Successful design and delivery of online professional development for teachers: A systematic review of the literature. *Computers & Education*, *166*, 104158. https://doi.org/10.1016/j.compedu.2021.104158

Butler, D. L., Lauscher, H. N., Jarvis-Selinger, S., & Beckingham, B. (2004). Collaboration and self-regulation in teachers' professional development. *Teaching and Teacher Education*, *20*(5), 435–455. https://doi.org/10.1016/j.tate.2004.04.003

Carrillo, C., & Flores, M. A. (2023). Online teaching and learning practices in teacher education after COVID-19: Lessons learnt from the literature. *Journal of Education for Teaching*, *49*(5), 869–881. https://doi.org/10.1080/02607476.2022.2153018

Chan, S., Maneewan, S., & Koul, R. (2023). Teacher educators' teaching styles: Relation with learning motivation and academic engagement in pre-service teachers. *Teaching in Higher Education*, *28*(8), 2044–2065. https://doi.org/10.1080/13562517.2021.1947226

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

Collins, L. J., & Liang, X. (2015). Examining high quality online teacher professional development: Teachers' voices. *International Journal of Teacher Leadership*, *6*(1), 18–34.

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., Saunders, L., & Coe, R. (2015). *Developing great teaching: Lessons from the international reviews into effective professional development*. Teacher Development Trust.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Dede, C., Jass Ketelhut, D., Whitehouse, P., Breit, L., & McCloskey, E. M. (2009). A research agenda for online teacher professional development. *Journal of Teacher Education*, *60*(1), 8–19. https://doi.org/10.1177/0022487108327554

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society, and Education*, *7*(3), 252–263. https://doi.org/10.25115/psye.v7i3.515

Didion, L., Toste, J. R., & Filderman, M. J. (2020). Teacher professional development and student reading achievement: A meta-analytic review of the effects. *Journal of Research on Educational Effectiveness*, *13*(1), 29–66. https://doi.org/10.1080/19345747.2019.1670884

Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, *61*(5), 713–740. https://doi.org/10.1177/0013164401615001

Fischer, C., Fishman, B., Dede, C., Eisenkraft, A., Frumin, K., Foster, B., Lawrenz, F., Levy, A. J., & McCoy, A. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform. *Teaching and Teacher Education*, *72*, 107–121. https://doi.org/10.1016/j.tate.2018.02.011

Fishman, B. (2016). Possible futures for online teacher professional development. In C. Dede, A. Eisenkraft, K. Frumin, & A. Hartley (Eds.), *Teacher learning in the digital age: Online professional development in STEM education* (pp. 13–30). Harvard University Press.

Fishman, B., Konstantopoulos, S., Kubitskey, B. W., Vath, R., Park, G., Johnson, H., & Edelson, D. C. (2013). Comparing the impact of online and face-to-face professional development in the context of curriculum implementation. *Journal of Teacher Education*, *64*(5), 426–438. https://doi.org/10.1177/0022487113494413

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Fütterer, T., Richter, E., & Richter, D. (2024). Teachers' engagement in online professional development - The interplay of online professional development quality and teacher motivation. *Zeitschrift für Erziehungswissenschaft*, *35*(4), 203–220. https://doi.org/10.1007/s11618-024-01241-8

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945. https://doi.org/10.3102/00028312038004915

Goodboy, A. K., & Martin, M. M. (2020). Omega over alpha for reliability estimation of unidimensional communication measures. *Annals of the International Communication Association*, *44*(4), 422–439. https://doi.org/10.1080/23808985.2020.1846135

Graham, J. W. (Ed.). (2012). *Missing data: Design and analysis*. Springer. https://doi.org/10.1007/978-1-4614-4018-5

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In M. Sehami, A. Fox, M. A. Hearst, & M. T. H. Chi (Eds.), *Proceedings of the First ACM Conference on Learning @ Scale Conference* (pp. 41–50). ACM. https://doi.org/10.1145/2556325.2566239

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But…. *Communication Methods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction*, *42*, 104–112. https://doi.org/10.1016/j.learninstruc.2016.01.007

Howard, N.-J. (2021). Barriers and drivers in online micro-course professional development: Navigating issues of teacher identity and agency. *Teaching and Teacher Education*, *105*, 1–12. https://doi.org/10.1016/j.tate.2021.103397

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hubers, M. D., Endedijk, M. D., & van Veen, K. (2020). Effective characteristics of professional development programs for science and technology education. *Professional Development in Education*, *48*(5), 827–846. https://doi.org/10.1080/19415257.2020.1752289

Jansen in de Wal, J., den Brok, P. J., Hooijer, J. G., Martens, R. L., & van den Beemt, A. (2014). Teachers' engagement in professional learning: Exploring motivational profiles. *Learning and Individual Differences*, *36*, 27–36. https://doi.org/10.1016/j.lindif.2014.08.001

Lefstein, A., Louie, N., Segal, A., & Becher, A. (2020). Taking stock of research on teacher collaborative discourse: Theory and method in a nascent field. *Teaching and Teacher Education*, *88*, 102954. https://doi.org/10.1016/j.tate.2019.102954

Lipowsky, F., & Rzejak, D. (2021). *Fortbildungen für Lehrpersonen wirksam gestalten. Ein praxisorientierter und forschungsgestützter Leitfaden [Designing effective professional development courses for teachers. A practice-oriented and research-based guide]*. Bertelsmann Stiftung. https://doi.org/10.11586/2020080

Lowes, S., & Lin, P. (2015). Learning to learn online: Using locus of control to help students become successful online learners. *Journal of Online Learning Research*, *1*(1), 17–48.

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research*, *44*(6), 764–802. https://doi.org/10.1080/00273170903333665

Martin, F., & Borup, J. (2022). Online learner engagement: Conceptual definitions, research themes, and supportive practices. *Educational Psychologist*, *57*(3), 162–177. https://doi.org/10.1080/00461520.2022.2089147

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781410601087

McDonough, D. (2014). Providing deep learning through active engagement of adult learners in blended courses. *Journal of Learning in Higher Education*, *10*(1), 9–16.

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

Meyer, A., Kleinknecht, M., & Richter, D. (2023). What makes online professional development effective? The effect of quality characteristics on teachers' satisfaction and changes in their professional practices. *Computers & Education*, *200*, 104805. https://doi.org/10.1016/j.compedu.2023.104805

Morin, A. J. S., Marsh, H. W., Nagengast, B., & Francesca Scalas, L. (2014). Doubly latent multilevel analyses of classroom climate: An illustration. *The Journal of Experimental Education*, *82*(2), 143–167. https://doi.org/10.1080/00220973.2013.769412

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (7th ed.). Muthén & Muthén.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90–100. https://doi.org/10.1016/S0022-2496(02)00028-7

Onrubia, J., Roca, B., & Minguela, M. (2022). Assisting teacher collaborative discourse in professional development: An analysis of a facilitator's discourse strategies. *Teaching and Teacher Education*, *113*, 103667. https://doi.org/10.1016/j.tate.2022.103667

Ornelles, C., Ray, A. B., & Wells, J. C. (2019). Designing online courses in teacher education to enhance adult learner engagement. *International Journal of Teaching and Learning in Higher Education*, *31*(3), 547–557.

Reeves, T. D., & Pedulla, J. J. (2011). Predictors of teacher satisfaction with online professional development: Evidence from the USA's e-learning for educators initiative. *Professional Development in Education*, *37*(4), 591–611. https://doi.org/10.1080/19415257.2011.553824

Richter, E., Marx, A., Huang, Y., & Richter, D. (2020). Time for professional learning: An empirical study about timing and duration of teacher training. *Zeitschrift für Erziehungswissenschaft*, *23*(1), 145–173. https://doi.org/10.1007/s11618-019-00924-x

Richter, E., & Richter, D. (2024). Measuring the quality of teacher professional development: A large-scale validation study of an 18-item instrument for daily use. *Studies in Educational Evaluation*, *81*, 101357. https://doi.org/10.1016/j.stueduc.2024.101357

Russell, M., Carey, R., Kleiman, G., & Venable, J. D. (2009). Face-to-face and online professional development for mathematics teachers: A comparative study. *Journal of Asynchronous Learning Networks*, *13*(2), 71–87. https://doi.org/10.24059/olj.v13i2.1669

Ryan, T. G. (2019). Naturalistic observation of engagement and disengagement within professional development in education. *International Online Journal of Education and Teaching*, *6*(1), 37–54.

Salta, K., Paschalidou, K., Tsetseri, M., & Koulougliotis, D. (2022). Shift from a traditional to a distance learning environment during the COVID-19 pandemic. *Science & Education*, *31*, 93–122. https://doi.org/10.1007/s11191-021-00234-x

Simonson, M. (1999). Equivalency theory and distance education. *TechTrends*, *43*, 5–8. https://doi.org/10.1007/BF02818157

Simonson, M., Schlosser, C., & Hanson, D. (1999). Theory and distance education: A new discussion. *American Journal of Distance Education*, *13*(1), 60–75. https://doi.org/10.1080/08923649909527014

Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., van Herwegen, J., & Anders, J. (2021). *What are the characteristics of teacher professional development that increase pupil achievement? A systematic review and meta-analysis*. Education Endowment Foundation.

van Nuland, S., Mandzuk, D., Tucker Petrick, K., & Cooper, T. (2020). COVID-19 and its effects on teacher education in Ontario: A complex adaptive systems perspective. *Journal of Education for Teaching*, *46*(4), 442–451. https://doi.org/10.1080/02607476.2020.1803050

Vivante, I., & Vedder-Weiss, D. (2023). Examining science teachers' engagement in professional development: A multimodal situated perspective. *Journal of Research in Science Teaching*, *60*(7), 1401–1430. https://doi.org/10.1002/tea.21836

Wynants, S., & Dennis, J. (2018). Professional development in an online context: Opportunities and challenges from the voices of college faculty. *Journal of Educators Online*, *15*(1), 1–13. https://doi.org/10.9743/jeo2018.15.1.2

Yoon, S. A., Miller, K., & Richman, T. (2020). Comparative study of high-quality professional development for high school biology in a face-to-face versus online delivery mode. *Educational Technology & Society*, *23*(3), 68–80.

APPENDIX A
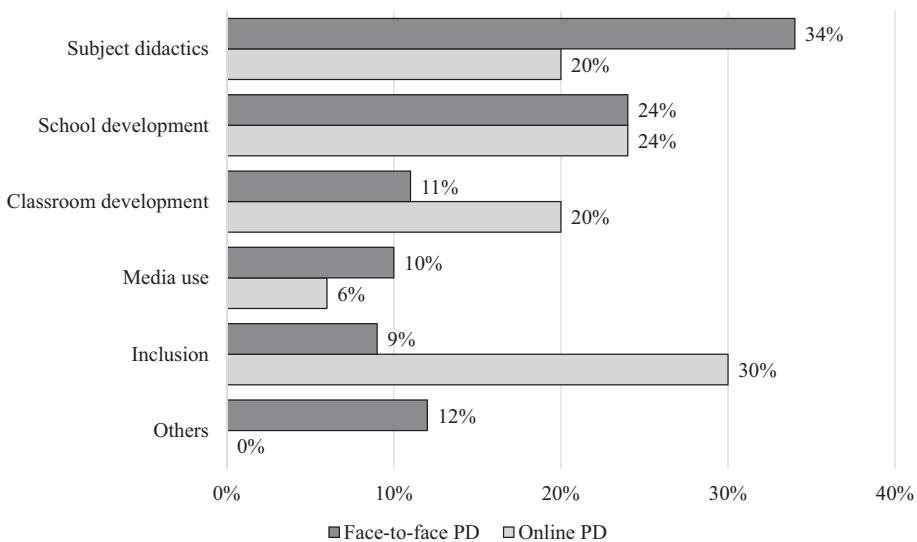
FIGURE A1 Relative frequencies of PD topics in face-to-face ($n = 137$) and online PD ($n = 54$). The category 'others' includes topics such as internal differentiation, performance diagnostics and violence prevention.

TABLE A1 Questionnaire for the survey of PD quality.

		Strongly disagree	Somewhat disagree	Somewhat agree	Strongly agree
1. How would you rate the clarity and structure of the event you attended?					
(a)	The goals of the event were clearly stated	○	○	○	○
(b)	The event content was explained in an understandable way	○	○	○	○
(c)	During the event, the time was used efficiently	○	○	○	○
(d)	The event schedule was clarified at the beginning	○	○	○	○
(e)	The individual components of the event built on each other in a meaningful way	○	○	○	○
2. How did the event relate to your professional practice?					
(f)	The discussion of the contents was based on real examples from professional practice	○	○	○	○
(g)	The aspects covered in the event were related to my current professional practice	○	○	○	○
(h)	The event addressed problems that I encounter in my everyday professional life	○	○	○	○
(i)	What I learned in the event I can apply in my professional practice	○	○	○	○
3. To what extent do you agree with the following statements about cognitive activation during the event?					
(j)	My prior knowledge was incorporated into the event	○	○	○	○
(k)	In the event, questions or tasks were asked that made me think	○	○	○	○
(l)	I was able to reflect on my professional tasks during the event	○	○	○	○
(m)	The feedback I received during the event provided me with concrete advice for my professional development	○	○	○	○
(n)	The event provided an opportunity to question previous routines in my work	○	○	○	○
(o)	During the event, I was able to think my way into the unknown	○	○	○	○
4. To what extent did the event provide opportunities for exchange with other participants?					
(p)	During the event, I had the opportunity to discuss the event content in depth with the other participants	○	○	○	○
(q)	The event provided an opportunity to work on tasks with other participants	○	○	○	○
(r)	During the event, work was facilitated in small groups	○	○	○	○
5. Finally, please share your overall assessment of the event					
(s)	I was able to comprehensively expand my competencies during the event	○	○	○	○
(t)	I will recommend the event to others	○	○	○	○

TABLE A2 Results of the multilevel regression analyses: PD format as predictor of PD quality without covariates (y-standardized regression coefficients).

Predictors	Clarity and structure Model 1a $\hat{\beta}$ (SE)	Cognitive activation Model 1b $\hat{\beta}$ (SE)	Collaboration Model 1c $\hat{\beta}$ (SE)	Practical relevance Model 1d $\hat{\beta}$ (SE)
Group level				
PD format (1 = online)	−0.30** (0.09)	−0.33** (0.09)	−0.49** (0.09)	−0.22* (0.10)
R^2 (between)	0.09	0.11	0.24	0.05
R^2 (within)	−	−	−	−
χ^2	77.76	83.70	17.71	12.86
df	18	28	4	10
p	<0.01	<0.01	<0.01	0.23
CFI	0.97	0.98	0.97	0.99
RMSEA	0.04	0.03	0.04	0.01
SRMR between	0.09	0.09	0.18	0.06
SRMR within	0.04	0.02	0.02	0.01

*$p < 0.05$; **$p < 0.01$.

TABLE A3 Results of the multilevel regression analyses: PD format as predictor of PD engagement without covariates (y-standardized regression coefficients).

Predictors	Behavioural engagement Model 2a $\hat{\beta}$ (SE)	Cognitive engagement Model 2b $\hat{\beta}$ (SE)	Affective engagement Model 2c $\hat{\beta}$ (SE)
Group level			
PD format (1 = online)	−0.43** (0.08)	−0.47** (0.09)	−0.28** (0.08)
R^2 (between)	0.19	0.22	0.08
R^2 (within)	−	−	−
χ^2	88.44	<0.01	78.02
df	10	<0.01	10
p	<0.01	1.00	<0.01
CFI	0.96	1.00	0.98
RMSEA	0.06	<0.01	0.06
SRMR between	0.06	<0.01	0.05
SRMR within	0.04	<0.01	0.03

**$p < 0.01$.