



## Teachers' participation in evaluating a web-based tool to monitor intervention fidelity

Jeanette Halvorsen, Sigrun K. Ertesvåg & Pål Roland

To cite this article: Jeanette Halvorsen, Sigrun K. Ertesvåg & Pål Roland (2023): Teachers' participation in evaluating a web-based tool to monitor intervention fidelity, Educational Research, DOI: [10.1080/00131881.2023.2223594](https://doi.org/10.1080/00131881.2023.2223594)

To link to this article: <https://doi.org/10.1080/00131881.2023.2223594>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 20 Jun 2023.



Submit your article to this journal [↗](#)






View related articles [↗](#)



View Crossmark data [↗](#)

## Teachers' participation in evaluating a web-based tool to monitor intervention fidelity

Jeanette Halvorsen , Sigrun K. Ertesvåg  and Pål Roland 

Norwegian Centre for Learning Environment and Behavioural Research in Education, Faculty of Arts and Education, University of Stavanger, Stavanger, Norway

### ABSTRACT

**Background:** When educational interventions are carried out, it is important that they are undertaken in a way that is aligned with the intervention plan: in other words, that they demonstrate fidelity to the intervention. A significant research issue is how fidelity can be monitored in a time-efficient and cost-effective way in classrooms and whether technology could help to provide innovative solutions in this regard.

**Purpose:** Through collaboration with teachers, this exploratory study sought to ascertain the usability of a web-based fidelity application (WFA). The WFA was being developed as a checklist tool to help teachers monitor the implementation fidelity of a social and emotional learning intervention for 14- to 15-year-old students in Norway.

**Method:** For this qualitative study, data were collected at two time points: (i) through a focus group interview with six teachers who had piloted the WFA prior to the initiation of the intervention; and (ii) via individual interviews with five teachers in the intervention group who had used the WFA during the implementation process. The data were analysed thematically.

**Findings:** According to the analysis, the teachers considered that the WFA's features (e.g. layout and registration process) could help support the ease and efficiency of fidelity reporting. They felt that it provided a highly recognisable link with the intervention material. In addition, the teachers provided ideas for further development and potential improvements. In all, the WFA was perceived as having high usability, suggesting its potential value as a useful tool for the collection of fidelity data.

**Conclusion:** This paper highlights the crucial role of teacher participation and the importance of fidelity data in the conduct of educational interventions. It draws attention to the need for user-friendly tools to support teachers to monitor fidelity in ways that do not involve high time and cost burdens. Similar WFAs could be of potential use in many different kinds of educational interventions in classroom settings internationally.

### ARTICLE HISTORY

Received 12 September 2022

Accepted 7 June 2023

### KEYWORDS

Intervention fidelity; educational intervention; implementation science; social and emotional learning; usability research; educational web-based applications

**CONTACT** Jeanette Halvorsen  [jeanette.halvorsen@uis.no](mailto:jeanette.halvorsen@uis.no)

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

## Introduction

The development and use of evidence-based educational interventions in the classroom is vital to supporting high-quality teaching and learning (Moir 2018). When interventions are carried out, whether by researchers, educators and/or other educational professionals, it is important that they are delivered in a way that is consistent with the intervention plan – in other words, that they demonstrate fidelity to the intervention. Thus, operationalised as ‘the extent to which implementers adhere to the intended treatment model’ (Humphrey et al. 2016b, 6), fidelity data provide information on the degree to which deliverers implement the structure and sequence of activities as planned by the intervention developer(s) (Humphrey et al. 2016a). One way fidelity can be monitored is via observation. However, this is time-consuming and may ultimately prove cost-ineffective, as observations are often restricted by associated costs (Berkel et al. 2011; Blase et al. 2012; Durlak and DuPre 2008), which may result in limited observations involving only a subsample of the intervention.

Therefore, a significant research issue is how fidelity can be monitored in a more time-efficient and cost-effective manner that is feasible in complex delivery settings such as classrooms, together with the extent to which technological applications may help provide innovative solutions. In the field of implementation science (Albers, Shlonsky, and Mildon 2020), there is a need for greater focus on evaluations of fidelity in educational interventions (Humphrey et al. 2021) and explorations of the usability of educational technology (Lu et al. 2022) in this regard. The study reported in this paper sought to contribute to this field of enquiry by evaluating the usability of a web-based fidelity application (WFA) for teachers, which was designed with ease of use and efficiency in mind, and developed with the participation of teachers. The WFA is a tool for monitoring fidelity in a social and emotional learning (SEL) programme (known as the Resilient Intervention (RI)), which was implemented among 14- to 15-year-old students in Norway. It is hoped that the study findings will have wider applicability. For instance, we anticipate they may be of interest to those developing educational interventions and working in a participatory way with teachers to support implementation quality in other settings internationally. In advance of explaining our research further, however, we situate our work within its theoretical framework and study context.

## Background

### *Conceptual background*

#### *Fidelity*

Along with dosage, quality, reach, responsiveness, programme differentiation, monitoring of control/comparison groups and adaptation, fidelity/adherence is an aspect of implementation and process evaluation (IPE) (Humphrey et al. 2016b). IPE enables insight into the impact mechanisms of educational interventions through theoretical, methodological and analytical tools (Humphrey et al. 2016a). The provision of manuals, guidelines, training and feedback may be used to optimise implementation fidelity (Durlak and DuPre 2008). A positive outcome in implementing a new programme is contingent on the programme being executed as prescribed (Gage et al. 2020). Whilst

there are no established criteria for the degree of fidelity necessary for an intervention to be effective (Gage et al. 2020), the more complex the activities in the intervention, the lower the fidelity of the implementation of these activities is expected to be (Gresham 2017).

Crucially, if fidelity is not evaluated, it is not possible to determine whether a lack of impact may be due to poor implementation or inadequacies inherent in the programme itself (Carroll et al. 2007). When programme fidelity is poor and facilitators do not deliver core components in line with the manuals, guidelines and training that have been provided, the outcome quality is likely to diminish (Berkel et al. 2011).

### Usability

Usability of educational and learning technology is not clearly defined (Lu et al. 2022). In our study, it was understood as ‘the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use’ (The International Organization for Standardization 2018). To achieve the quality standard needed, it was determined that the WFA that was developed in our study should be easy to use by the intended users, reliable, structured and valid (Ibrahim and Sidani 2015). It should also be effective, leave no room for errors and provide ease of learning (Rusu et al. 2015). Schmidt and Huang (2022) state that user-friendly systems tend to be more efficient, satisfying and effective, while Fernandez, Insfran, and Abrahão (2011) report that high usability is the most important quality factor for web applications.

In the development of the WFA in our study, five categories were applied to ensure an in-depth evaluation and assessment of high-quality usability. These comprised *effectiveness*, which supports users in completing actions accurately; *efficiency*, which ensures that users can perform tasks quickly through the easiest process; *error tolerance*, which indicates common errors that users make and how easily users can recover from those errors; *ease of learning*, which provides information on how new users can easily accomplish goals and, lastly, *engagement*, which provides information on the extent to which users find it (in our case, the WFA) pleasant to use and appropriate for its industry/topic (Fernandez, Insfran, and Abrahão 2011; The International Organization for Standardization 2018; Rusu et al. 2015).

### Study context

As mentioned earlier, in our study, the WFA was used to monitor fidelity in the Resilient Intervention (RI), which was a social and emotional learning programme. This took the form of a randomised controlled trial (RCT) where the intervention was a comprehensive SEL curriculum (hereafter ROBUST) and professional development programme for teachers. It was implemented in 24 9th-grade middle schools (43 classes of 14- to 15-year-old students) in five Norwegian municipalities. The RI aimed to enhance students’ wellbeing and motivation, reduce emotional distress and loneliness, and improve academic outcomes by fostering competencies in the five core components presented below. Teachers received digital training on methods for implementing core components, including a detailed plan for scheduling. Access to supervision, a resource book describing the content of each lesson and online resources (including fully developed electronic

presentations, podcasts and student materials) were available. A step-by-step instructional video was created to illustrate all components of utilisation.

The proposed WFA being developed for use by teachers was a web application: in other words, a type of software that allows users to interact with a remote server through a web browser interface (The International Organization for Standardization 2017). The function of the Resilient WFA was to provide teachers with the ROBUST Fidelity Checklist (Ertesvåg et al. 2020) for them to use. As part of the development of the WFA, this study focused on teachers' early utilisation behaviour to detect problems and provide opportunities for improvement (Lu et al. 2022). The ROBUST intervention itself consisted of five core components: social relations, mindfulness, problem solving, emotions and growth mindset. Each core component contained five lessons, with five to seven activities in each, for 25 total ROBUST lessons. The individual lesson times were 45 min, on average. Three lessons from each core component were designed to be delivered sequentially, followed by the last two lessons from each of the components.

Delivery of the core components was monitored through the teachers' registration of completed lessons. The WFA registration process was designed to be effective and engaging, with each registration taking approximately 2 min. Teachers had to follow a link directly to the registration screen, which was designed with the same recognisable layout as the resource book, where they had to add their unique information (e.g. teacher ID). To complete registration, the teacher was required to click on the specific lesson and add the time, date and number of students present. When all activities in a lesson had been scored, that lesson was registered as completed. The activities in each lesson followed the planned order of delivery. The scoring format, presented in different colours, was *everything*, *something* or *nothing*. An overview of progress was available to the teachers on the WFA.

### **Purpose**

With the above context and research setting in mind, the overall purpose of the current study was to gain insight into the usability of the web-based fidelity application, from the teachers' perspectives. The research question was 'How do teachers evaluate the usability of the web-based fidelity application?'

### **Methods**

A qualitative research methodology (Denzin and Lincoln 2018) and exploratory design (Miles, Huberman, and Saldaña 2020) was selected as most suitable for the study's purpose, as it was necessary to collect rich data from participants who had experienced the WFA and perform an in-depth analysis on the data collected. Two types of data collection were undertaken, at different time points: (i) a focus group interview with teachers who had piloted the WFA and (ii) individual interviews with teachers who were implementing the full version of ROBUST. Data were collected across the initiation and start-up phases of the RI. As a consequence of the COVID-19 pandemic, the interviews were conducted using a videoconferencing platform, thus allowing the inclusion of participants from all over Norway. Disadvantages of conducting interviews virtually include the risk of technical difficulties and the interviewer's lack of control over the

environment during the interview (Brinkmann 2018; Krueger and Casey 2015). To mitigate these issues as far as possible, participants received clear, written instructions about participation in advance of the interview.

Data from the focus group interview, and the individual interviews conducted at different time points throughout the implementation process, were combined. It was anticipated that the two data sources, in combination, could provide complementary perspectives and potentially yield rich data, as the two forms of interviews would offer different viewpoints in terms of participants' interaction with the WFA (Blossing, Roland, and Sølvik 2019). Specifically, whilst the individual interviews would offer varied and unique perceptions, a group perspective may contribute to reflection development and broader insights at the collective level through interactions in the group, as participants would be able to comment on suggestions from other participants (Stewart and Shamdasani 2015).

### *Ethical considerations*

The study was registered with the Norwegian Social Science Data Services (number 981143) and evaluated to be in accordance with the Norwegian Privacy Act (Norwegian Data Inspectorate 2022). Participants were informed that their participation was voluntary and that they could withdraw from the study at any time without consequences; they provided written consent (The Norwegian National Research Ethics Committees 2022). The participants' identities were anonymised using codes, and the recorded data were stored separately. A secure server was used for the WFA, and participants used an anonymised ID with a secure password each time at login.

### *Data collection*

#### *Focus group interview*

A group of six teachers piloted the WFA and tested five carefully selected lessons, with one lesson from each core component, in their own classrooms. This pilot study was conducted prior to the implementation of the RI, and participants signed a nondisclosure agreement. The lessons that were tested out in the pilot were selected in cooperation with the programme developers, researchers and implementers of the RI. Participants had been recruited through the official social media channels of the Norwegian Centre for Learning Environment and Behavioural Research in Education. The selection criterion for participation was teaching in 9th grade at middle schools in municipalities not taking part in the RI.

A focus group interview (Krueger and Casey 2015) was conducted with these six teachers (i.e. 'teacher pilot participants', or 'TPP'). These teachers were not part of the intervention group, and the interview was conducted prior to the initiation of the RI. The participants were in majority female and had an average age of 38, an educational background in secondary education and an average of 11 years of teaching experience. Computers and mobile phones had both been used to access the WFA, and the teachers had various experiences of using web applications in teaching. The focus group session was carried out in May 2021. Initially, a pilot interview was undertaken to confirm the time frame and the quality of both the video and audio when recorded. Focusing on the

theoretical framework, a semi-structured interview guide was developed, containing open-ended questions that allowed for follow-up questions (Brinkmann 2018; Miles, Huberman, and Saldaña 2020). The pilot interview contributed necessary information on how to further develop the interview guide to best fit the exploratory purpose of the research (Denzin and Lincoln 2018).

The use of a moderating team, with clear roles and responsibilities for the moderator and assistant moderator, provided a mutual basis for discussion when analysing the interview data at a later stage (Krueger and Casey 2015; Malterud 2012). The moderator distributed the questions. To prevent the risk of a single participant's voice influencing the group, the assistant moderator ensured that all participants had the opportunity to convey their own experiences through equal participation. This made it possible to go deeply into particular subjects in group discussions when needed (Stewart and Shamdasani 2015). The duration of the focus group interview was 1.5 h, and the language of the interview was Norwegian.

### *Individual interviews*

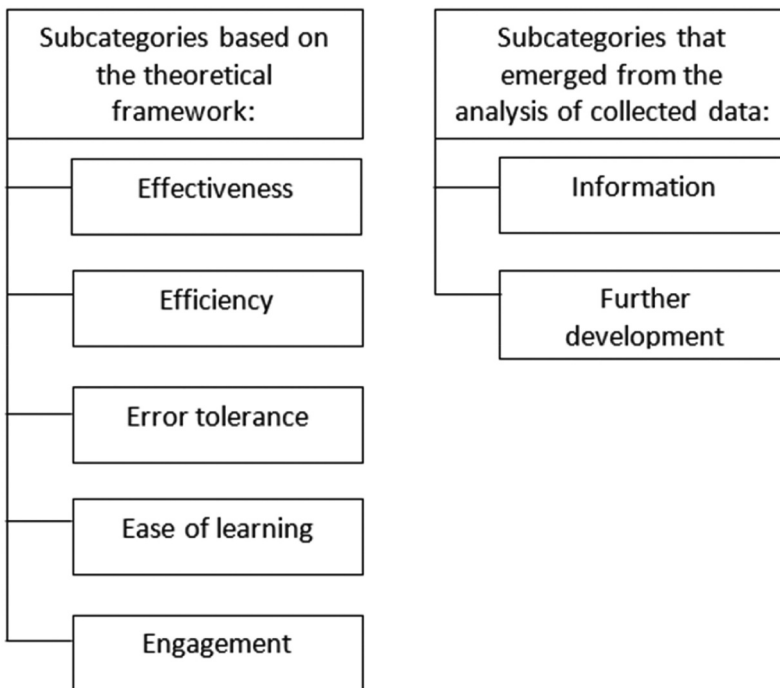
Individual interviews were then carried out with a different group of teachers. These were teachers who had implemented the full version of the intervention (i.e. teachers who were implementing 'ROBUST' (i.e. the SEL curriculum); or 'TIR' participants). These participants were recruited from the RI group to represent variations in WFA reporting. They were teachers who had utilised the WFA in most lessons, some lessons or no lessons from the planned scheduled programme in the RI. Five teachers participated in the individual interviews. They were in majority female and had an average age of 36 years old, an educational background in secondary education and an average of 8 years of teaching experience. As with the TPP group, computers and mobile phones had both been used to access the WFA, and participants had various experiences of using web applications in teaching.

These individual interviews took place during the period from November 2021 to March 2022. They were semi-structured interviews which were carried out using the same interview guide as had been used for the focus group interview (Brinkmann 2018). The participant and moderator were present during individual interviews. The interviews were conducted in Norwegian and lasted between 16 and 53 min based on each participant's experience with the WFA.

### *Data analysis*

The study employed an interactive approach (Miles, Huberman, and Saldaña 2020) to analyse the data in three stages: (a) data condensation, (b) data display and (c) drawing conclusions. During the data condensation process, the most important information was highlighted by means of simplifying, transforming and abstracting the data at hand (Malterud 2012). A deductive analysis procedure (Hsieh and Shannon 2005) with an exploratory approach (Denzin and Lincoln 2018) was selected. This allowed us to make sense of the data closely linked to the theoretical framework. Below, we describe the stages in the analysis in more detail.

First, the focus group and individual interviews were transcribed (Brinkmann 2018), resulting in a total of 125 pages. The first author coded the transcripts into separate codes



**Figure 1.** Subcategories from the analysis.

and nodes, using qualitative analysis software as a tool. Based on the theoretical framework and emergent categories, the codes were discussed between co-authors. When combining the focus group interview and the individual interviews, the point of saturation was reached when, during coding, no new nodes occurred because of informational redundancy (Saunders et al. 2018). Through this overview, the material became clearer and more accessible. Second, a display of the data was made, presenting an organised, compressed and accessible visualisation. Third, compressed data from the interviews were compared, and the main patterns were extracted, as presented in Figure 1. This contributed to the deductive, in-depth analysis of the categories identified and developed from the theoretical framework.

Subsequently, a third category, presented in Figure 2, was identified from the emerging material. As the emerging data showed differing patterns between the focus group interview and the individual interviews, merging patterns were further compared to demonstrate how any contrasting findings complemented each other (Hsieh and Shannon 2005). Emerging from the two different types of interviews conducted at different time points, the findings from the analysis may complement each other and thus yield richer data. The assistant moderator contributed to the data analysis by reading all interview transcripts and taking part in discussions. Aspects of the study were discussed in internal research groups.

## Findings

The analysis of focus group and individual interview data allowed us to explore the participants' views about the usability of the WFA, thereby affording insight into our research question. Overall, the analysis led to the subcategories presented in Figure 1,





an individual interview participant suggested that there should be a link to the WFA on the last slide of each presentation leading directly to registration.

The participants reported in general that they felt that the WFA functioned efficiently. This was exemplified by a focus group participant observing that the registration was well adapted to the idea of reporting whether the lesson was delivered as planned. However, an individual interview participant commented that registering accurate information presupposed familiarity with the ROBUST curriculum and what needed to be done in specific lessons. This contradicted the general focus group view that registration was self-explanatory. In addition, several participants used terms such as *registration tool* when they discussed the efficiency of the WFA, illustrated here by an individual interview participant:

The process of registering what you have done is very efficient; I have used it [the WFA] as a registration tool.

The focus group participants reported that they required 1–2 min to complete the report for each of the five ROBUST lessons that they implemented, further indicating efficient registration.

With respect to error tolerance, some of the individually interviewed participants pointed to aspects of the WFA that required precision to record accurate information. One example provided was the variability of sensitivity in setting the date and time. Another expressed some confusion regarding the WFA not being downloadable and the fact that it was still referred to as an application in the project rather than a web application. A further participant stated that no errors preventing registration in the WFA were detected during use. These examples did not change the general view that the WFA was effective and efficient in supporting teacher registration of fidelity in an easy-to-use format.

### **Overarching category 2: engagement strategies**

This category consisted of the subcategories *ease of learning* and *engagement* (see [Figure 2](#)). For the WFA to be easy to learn and engage users, users should find it easy to accomplish goals while utilising the tool; furthermore, it should be pleasant to use and appropriate for its topic. When it came to ease of learning, the majority of participants perceived it as an intuitive, logical tool that provided a highly recognisable link between the ROBUST resource book and content in the lessons, allowing them to easily accomplish goals. For example, a focus group participant observed:

I like that the different exercises follow each other in the same order as in the resource book; it provides a good overview. I also think the front page is nice, where you have a full overview of topics and see a progression on how much you have done.

Whilst the focus group participants reported gradual familiarisation in the learning process, indicating that registration was more demanding the first time, the individually interviewed participants tended to feel that registering was easy from the beginning. This finding might relate to the fact that the focus group participants, who had been involved in the pilot, did not receive training prior to utilisation, whereas the individually interviewed participants, who had been delivering the intervention itself, did. Following the

link to the WFA each time was regarded as impractical by some of the focus group, possibly contradicting the ease of the learning process according to some of the individually interviewed participants. In general, several participants stated that saving the WFA as a shortcut or icon on the computer desktop solved this problem, enabling them to easily find it directly after completing lessons.

In relation to the topic of engagement, the main tendency among the two teacher groups was that the WFA initiated engagement and that it was pleasant to use. An individual interview participant compared their experiences with the WFA to those with other applications used in a teaching context, remarking as follows:

It looks appealing, and it is engaging in such a way that it makes you want to click on it.

The teachers who had been involved in the pilot commented that they found the WFA engaging to use, noting that it had the potential to be utilised to help track which lessons had been implemented. The individually interviewed teachers agreed with this to some extent, focusing on feedback and how the WFA was perceived as engaging and appropriate for its intended use. For example, an individual interview participant compared the colour scheme to a feeling of 'reward and punishment', based on the colours that appeared on the screen when marking the options *everything* (green), *something* (orange) or *nothing* (white). This could be regarded as an indication that the WFA was not as pleasant as it could be to use, thus contradicting the main view. This participant further explained that, for a teacher keen to comply with the programme, this feedback would be engaging because the participant would want to do *everything* in the ROBUST curriculum in order to build up as many green markings as possible. The individually interviewed participants further viewed the WFA as appropriate for its topic, frequently mentioning that it was well adapted to its purpose when aiming to collect information on implementation in the classrooms.

### **Overarching category 3: further development**

As illustrated in [Figure 2](#), this overarching category consisted of the two supplemental subcategories of *information* and *further development* that emerged from the in-depth data analysis. The overarching category contributed beyond the theory-driven categories, yielding some valuable insights that were needed to gain a broader understanding of how useable the WFA was.

In terms of information, there was a recurring interest among the teachers who had been involved in the pilot study for more information and more training prior to utilisation of the WFA. For instance, one focus group participant suggested that there should be an introduction to the utilisation of the WFA, while another focus group participant pointed out that the teacher training should have been different:

We should have received more training on how to use the app and a short introduction to its potential, both in ROBUST and in teaching. That would have been useful.

The focus group participant further described a feeling of not being able to give the desired feedback, whilst other participants wondered whether the information they provided was sufficient for the project. Another focus group participant followed up on the statement by observing that registration did not allow for complementary

information and that the possibility of providing broader feedback in the form of text in an additional comment section might be useful. However, the idea that the efficiency of reporting in the WFA allowed for spending more time with students in the classroom was raised by a further focus group participant. Interestingly, the individually interviewed participants also suggested that a method for providing more in-depth information beyond responding with *everything*, *something* or *nothing* in reports would be helpful. For instance, one participant reflected on this issue and what it might require:

I wanted an opportunity to explain why I did not have time to do it [registering in the WFA]. At the same time, the intention is that it should be a fast process that does not require much work. If you must write about or start justifying why you have done 'nothing', then it will be just that - a lot of work.

In addition, several of the teachers who had been involved in the pilot suggested that an instructional video showing that using the WFA was intended to be as straightforward as perceived would be beneficial. Furthermore, some proposed that the ROBUST material (e.g. descriptions of various lessons and exercises) could have been placed directly in the WFA to concretise and guide the registration process. As an example, a focus group participant explained that if everything could be gathered in one place, with an overview of lessons and all material needed for completing them, it would be easier to implement in the classroom. Another focus group participant, who had only had access to a few lessons, suggested that, with guidelines within the WFA itself, the application might have even wider potential.

Directly indicating further development, the individually interviewed participants provided supplemental, in-depth information on topics raised by the pilot group participants. Specifically, the teachers who had participated in the full implementation generally supported the pilot study participants' experience of the WFA being time-efficient and easy to use. Another aspect of further development was mentioned by an individual interview participant reflecting on the importance of registering directly after each completed lesson:

Since I did not always remember to fill in directly after each completed lesson, it became difficult to remember everything. I was afraid that I had forgotten something and that my reporting was inaccurate.

As a result of not remembering or not having time to register in the WFA immediately after each completed lesson, one participant suggested that a reminder received via a notification might be helpful. However, others countered that such a notification could be bothersome.

As already touched upon, the desire to register broader feedback in an additional comment section was reported by the focus group and individually interviewed participants alike. A focus group participant suggested the practicality of noting information one would want to remember, review or assess more thoroughly at a later stage. It was felt that the wish to give complementary feedback was prominent but also had two sides: the need to report more information would engage and motivate, but, at the same time, the possibility to do so would demand more work from the participants. This is aptly illustrated by an individual interview participant's comment:

There is little room for complementary feedback, but I do not think I would have time to give long, complex, complementary feedback either.

Ultimately, it seemed that according to some of the participants' perceptions, the provision of an additional comment section or providing complementary feedback might make the WFA appear more complicated and less convenient, thus negatively affecting usability.

In summary, the participants, overall, tended to describe the WFA in favourable terms. They considered that it had a self-explanatory layout with an effective process of registration and viewed it as an intuitive, logical tool which was able to provide a highly recognisable link with the intervention material. In addition, a recurring interest in receiving more information and training prior to utilisation was noted. The analysis of the participants' perceptions suggested, encouragingly, that the WFA had the potential to be utilised to collect fidelity data in a way that was time-efficient and cost-effective.

## Discussion

Through collaboration with teachers, who variously piloted the WFA, used the WFA as part of an intervention and took part in the focus group or individual interviews, we were able to progress thinking about the WFA's usability and better understand how it could be further developed. In this section, we return to the overarching categories (*Structural requirements*, *Engagement strategies* and *Further development*) and consider the implications of our findings in relation to relevant literature.

### *Structural requirements: structure matters*

It was evident from the findings that teachers from the focus group and those who were individually interviewed alike considered the WFA to have high usability. This suggests that, for these users at least, the WFA had reached the quality factor necessary for a web application that allows users to interact over a remote server through a web browser (Fernandez, Insfran, and Abrahão 2011; The International Organization for Standardization 2017, 2018). In all, it was clear from analysis of the participants' feedback that the WFA was in line with the requirements highlighted in the research literature (Ibrahim and Sidani 2015; Rusu et al. 2015; Schmidt and Huang 2022). Furthermore, the data analysis indicated that the participants were enthusiastic about utilising an innovative, easy-to-use tool to monitor their own or the class's process(es) to keep track and control what had been implemented or not.

This approach could be transferable to areas of classroom practice to support high-quality teaching and learning (Moir 2018) other than research uses. For example, participants in both the focus group interview and the individual interviews suggested that the WFA should be utilised to register student progress in other subjects, thus making clear the potential of the WFA for further development in educational technology research (Lu et al. 2022). Possible benefits would be relevant to smaller projects and larger, more resource-demanding projects, too. The total number of lessons to be implemented in the ROBUST intervention among the 43 classes was 1075, meaning that observation would have been resource intensive, time-consuming and cost ineffective. Whilst observation is

widely regarded as the gold standard for collecting fidelity data, and self-reporting may be viewed as less reliable, the WFA could be considered a valuable tool for registering fidelity data where observation is not possible or feasible, due to its time-effective, cost-effective aspects (Blase et al. 2012) and user-friendly features, consistent with those suggested in the research literature (Schmidt and Huang 2022).

It is noteworthy that the views of the focus group and individually interviewed participants were not always in alignment, possibly because they experienced the WFA in different contexts and at different time points (i.e. the pilot study in the case of the focus group participants and the intervention itself in the case of the individually interviewed participants). Nonetheless, most participants felt that the WFA provided a highly recognisable link between the ROBUST resource book and the content in the lessons. The ROBUST Fidelity Checklist (Ertesvåg et al. 2020) reflected the core components and change mechanisms as described in the resource book. Since the RI also offered guidelines, digital resources, training and supervision to teachers in the intervention group, it can be argued that ROBUST was described in such detail that, in accordance with Durlak and DuPre (2008), it provided grounding for implementation with high fidelity. Thus, the WFA can be thought of as a tool that helped lay a foundation for teachers to deliver core components clearly and comprehensibly when executing ROBUST as prescribed for a positive outcome (Gage et al. 2020) and to avoid a decrease in the quality of outcome implementation (Berkel et al. 2011). This suggests, more broadly, that the WFA's potential for future utilisation lies in its adaptable structure, layout and regulation of content.

The different levels of training, information and access to material would have undoubtedly affected how the two teacher groups evaluated the WFA. For example, we found that the teachers who used the WFA in the pilot study perceived the speed of registering in the WFA as particularly valuable, whereas the teachers who used the WFA in the intervention itself especially welcomed room for complementary feedback. Such differences in emphasis might be related to differing levels of lessons delivered from the ROBUST curriculum, where a higher level of delivery might be connected to the need for broader feedback. As an example, the teachers who used the WFA in the intervention itself requested a way to give more in-depth information, at the same time recognising that this would require more work. As observed by Gresham (2017), it can be the case that the more complex the activities in an intervention are, the lower one can expect the fidelity of implementation of those activities to be. It is also evident that the complexity of an intervention is likely to increase the difficulty of monitoring fidelity.

Since the degree of fidelity necessary for an intervention to be effective is far from a straightforward question (Gage et al. 2020), monitoring fidelity early in the implementation process can be crucial (Humphrey et al. 2016b). Overall, the teachers' feedback in our study reflects that structure matters: in terms of the significance of the structure of the WFA as a registration tool and its role in monitoring implementation fidelity.

### *Engagement strategies: easy to learn, engaging to use*

In general, our study participants found the WFA engaging for its intended use and well adapted to its purpose of research and the collection of information on implementation fidelity in the classroom. This suggests that learning to use the WFA was easy, which is necessary if it is to be perceived as engaging (The

International Organization for Standardization 2018; Rusu et al. 2015). The findings gave insight into how participants tended to perceive that the WFA's appearance, layout and overview supported the process of registering. It appeared that the colour scheme used for the registration options (i.e. *everything* (green), *something* (orange) or *nothing* (white)) worked as a motivating factor for most participants: green markings indicated positive associations in terms of completing lessons. However, it is interesting to reflect on the implications of an aforementioned comparison of the colour scheme to feelings of reward/punishment. It draws attention to the notion that some aspects of the theoretical framework, including engagement, tend by their nature to be more subjective than others, such as efficiency. In such areas, where views may vary more widely, adapting the content in the WFA to optimise teachers' experience might prove particularly challenging and must, of course, be considered in relation to the overall usability of the WFA. In the case of the colour scheme, it is possible, too, that some perceptions might derive from the participants' reported desire to complete all lessons, and the fact that they were prevented from achieving this goal by lack of time, or changes disrupting the planned lessons. Being prevented from completing a lesson might contribute to a feeling of disappointment, which in turn may be exacerbated by the white field appearing as one presses the *nothing* option.

In sum, the findings related to engagement suggest that the WFA was regarded, in the main, as a practical, easy-to-use tool. The more engaging the WFA is to use, the more loyal implementers might be in utilising it, possibly contributing to higher fidelity to programme implementation.

### **Further development: potential for improvement**

In the study, participants were keen to gain additional information about how to use the WFA and to have more training prior to its use. Whilst the desire for a more thorough introduction via an instructional video was often noted by the pilot study teachers, this need was not mentioned as often by the teachers carrying out the intervention, suggesting that being provided with the necessary information and training prior to utilisation was associated with how convenient the WFA was perceived to be. As a relatively young research field, implementation science needs convenient methods for monitoring and evaluating fidelity in educational interventions (Albers, Shlonsky, and Mildon 2020; Humphrey et al. 2021).

The importance of registering on the WFA directly after each completed lesson is another reason why the WFA needs to be convenient to use. Although some participants repeatedly proposed that a reminder or notification to register in the WFA directly after each completed lesson would be helpful, others indicated that this type of reminder could be bothersome. The possibility of errors and inaccuracies in registration increase when registering in retrospect, which highlights the importance of registering directly after the delivery of a lesson. Inaccuracies in collected data can have significant implications: for example, an inconsistency in the number of students registered and students present in specific lessons might affect the outcome of an intervention. Not accounting for implementation variability can lead to biased data and may make it difficult to gauge the true potential of preventative interventions (Humphrey et al. 2021). With this in mind,

a carefully designed notification to remind teachers to register directly after each completed lesson might be a helpful additional feature.

It was evident that the pilot study teachers felt that the SEL ROBUST curriculum would be easier to implement if descriptions of the various lessons and exercises, with guidelines, were available directly in the WFA to help concretise and guide the registration process. The substantive point here resonates with Durlak and DuPre (2008) and Humphrey et al. (2016b), who observe that the chance of an intervention being implemented as planned increases as teachers have a clearer perception of what to do and how to do it. This underlines the importance of obtaining an understanding of how and why an intervention works or does not work (Carroll et al. 2007) by identifying problems in usability (Lu et al. 2022), faults or other factors preventing implementation with high fidelity early in the process (Humphrey et al. 2016a, 2016b).

Tensions within the framework were clearly illustrated by the lack of participant consensus regarding the (in)ability to give complementary feedback in the WFA. The trade-offs between various aspects of effectiveness and efficiency, illustrated by differing viewpoints regarding the (in)ability to register complementary feedback, were prominent in the findings. One consideration is that, as the WFA was developed for quantitative purposes and intended to be time-effective, the opportunity to give feedback would make it more qualitative in nature and time-consuming, thus conflicting with its overall purpose. In all, the WFA was designed to enable the monitoring of fidelity by providing information on the extent to which the teachers adhered to both the structure and sequence (Humphrey et al. 2016a) of the lessons in ROBUST, without it being a time-inefficient and cost-ineffective process.

### *Methodological considerations and limitations*

Collecting data from the two different teacher groups at two different time periods was advantageous, as it allowed access to participants' experiences before and after the initiation of the intervention and afforded insight into the utilisation of the WFA at different stages. However, it is important, too, to recognise that the design of the study (i.e. use of a focus group at one stage and individual interviews at the other) may have affected the ways that participants responded. In addition, aspects related to the COVID-19 pandemic complicated the recruitment process. In terms of the sample, the participants in both teacher groups were all under 40 years of age, and thus, a possible limitation is the exclusion of older teachers from the sample, who may have brought different perspectives and viewpoints. Given our aim of exploring the usability of one specific WFA, generalisation from this small, exploratory study is not intended. Nevertheless, aspects of the findings may be relevant for the assessment of similar WFAs.

### **Conclusion**

In interventions, high fidelity to the programme is of major importance. Teachers play valuable and crucial roles through their participation in educational interventions. Thus, it is vital that ways are found to provide user-friendly tools to support teachers to monitor fidelity in ways that do not involve high time and cost burdens. In this exploratory study, through teacher participation, we gained much-needed



insight into the usability of a web-based fidelity checklist application from the teachers' perspectives. In general, the teachers felt that the WFA was a self-explanatory tool that was easy to learn; it was considered to have high usability. Further longitudinal research on utilising the WFA will be needed to elucidate its full potential as a tool to monitor fidelity in educational research. Another important consideration is how the quality and richness of the information gathered from the cost-effective WFA approach might differ from that gathered by a direct observation approach.

The practice of developing and using evidence-based interventions in classrooms is inextricably linked to progress in teaching and learning to support all learners. It is hoped that the insights offered by this study will be of interest to other researchers and educational professionals internationally, especially those who are involved in the design and delivery of educational intervention studies.

## Acknowledgements

We are thankful to all the teachers who participated in the project in a time when COVID-19 brought many unforeseen challenges.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by the Research Council of Norway under grant number 299166.

## ORCID

Jeanette Halvorsen  <http://orcid.org/0000-0003-2684-8651>

Sigrun K. Ertesvåg  <http://orcid.org/0000-0001-7804-8571>

Pål Roland  <http://orcid.org/0000-0001-7200-2601>

## References

- Albers, B., A. Shlonsky, and R. Mildon. 2020. "En Route to Implementation Science 3.0." In *Implementation Science 3.0*, edited by B. Albers, A. Shlonsky, and R. Mildon, 1–38. Cham, Switzerland: Springer International Publishing.
- Berkel, C., A. M. Mauricio, E. Schoenfelder, and I. N. Sandler. 2011. "Putting the Pieces Together: An Integrated Model of Program Implementation." *Prevention Science* 12 (1): 23–33. <https://doi.org/10.1007/s11121-010-0186-1>.
- Blase, K. A., M. Van Dyke, D. Fixsen, and F. Bailey. 2012. "Implementation Science: Key Concepts, Themes and Evidence for Practitioners in Educational Psychology." In *Handbook of Implementation Science for Psychology in Education*, edited by B. Kelly and D. Perkins, 13–34. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781139013949.004>.
- Blossing, U., P. Roland, and R. M. Sølviik. 2019. "Capturing Sense-Made School Practice. The Activities of the Interviewer." *Scandinavian Journal of Educational Research* 63 (7): 1007–1021. <https://doi.org/10.1080/00313831.2018.1476404>.

- Brinkmann, S. 2018. "The Interview." In *The Sage Handbook of Qualitative Research*, edited by N. K. Denzin and Y. S. Lincoln, 984–1026. Thousand Oaks, CA: Sage.
- Carroll, C., M. Patterson, S. Wood, A. Booth, J. Rick, and S. Balain. 2007. "A Conceptual Framework for Implementation Fidelity." *Implementation Science* 2 (1): 40. <https://doi.org/10.1186/1748-5908-2-40>.
- Denzin, N. K., and Y. Lincoln. 2018. "Introduction: The Discipline and Practice of Qualitative Research." In *Handbook of Qualitative Research*, edited by N. K. Denzin and Y. S. Lincoln, 1–32. Thousand Oaks, CA: Sage.
- Durlak, J. A., and E. P. DuPre. 2008. "Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation." *American Journal of Community Psychology* 41 (3): 327–350. <https://doi.org/10.1007/s10464-008-9165-0>.
- Ertesvåg, S. K., K. Tharaldsen, L. Vestad, N. Grini, and E. Bru. 2020. *Robust Fidelity Check List*. Stavanger, Norway: Norwegian Center for Learning Environment and Behavioral Research in Education, University of Stavanger.
- Fernandez, A., E. Insfran, and S. Abrahão. 2011. "Usability Evaluation Methods for the Web: A Systematic Mapping Study." *Information and Software Technology* 53 (8): 789–817. <https://doi.org/10.1016/j.infsof.2011.02.007>.
- Gage, N., A. Macsuga, R. Detrich, and J. States. 2020. *Fidelity of Implementation in Educational Research and Practice*. Oakland, CA: The Wing Institute.
- Gresham, F. M. 2017. "Features of Fidelity in Schools and Classrooms: Constructs and Measurement." In *Treatment Fidelity in Students of Educational Intervention*, edited by G. Roberts, S. Vaughn, S. N. Beretvas, and V. Wong, 22–38. New York: Routledge.
- Hsieh, H. F., and S. E. Shannon. 2005. "Three Approaches to Qualitative Content Analysis." *Qualitative Health Research* 15 (9): 1277–1288. <https://doi.org/10.1177/1049732305276687>.
- Humphrey, N., A. Lendrum, E. Ashworth, K. Frearson, R. Buck, and K. Kerr. 2016a. *Implementation and Process Evaluation (IPE) for Interventions in Education Settings: A Synthesis of the Literature*. London: Education Endowment Foundation.
- Humphrey, N., A. Lendrum, E. Ashworth, K. Frearson, R. Buck, and K. Kerr. 2016b. *Implementation and Process Evaluation (IPE) for Interventions in Educational Settings: An Introductory Handbook*. London: Education Endowment Foundation.
- Humphrey, N., M. Panayiotou, A. Hennessey, and E. Ashworth. 2021. "Treatment Effect Modifiers in a Randomized Trial of the Good Behavior Game During Middle Childhood." *Journal of Consulting and Clinical Psychology* 89 (8): 668–681. <https://doi.org/10.1037/ccp0000673>.
- Ibrahim, S., and S. Sidani. 2015. "Fidelity of Intervention Implementation: A Review of Instruments." *Health* 7 (12): 1687–1695. <https://doi.org/10.4236/health.2015.712183>.
- Krueger, R. A., and M. Casey. 2015. *Focus Groups: A Practical Guide for Applied Research*. Los Angeles: Sage Publications.
- Lu, J., M. Schmidt, M. Lee, and R. Huang. 2022. "Usability Research in Educational Technology: A State-Of-The-Art Systematic Review." *Educational Technology Research & Development* 70 (6): 1951–1992. <https://doi.org/10.1007/s11423-022-10152-6>.
- Malterud, K. 2012. "Systematic Text Condensation: A Strategy for Qualitative Analysis." *Scandinavian Journal of Public Health* 40 (8): 795–805. <https://doi.org/10.1177/1403494812465030>.
- Miles, M. B., A. Huberman, and J. Saldaña. 2020. *Qualitative Data Analysis: A Methods Sourcebook*. Thousand Oaks, CA: Sage Publications.
- Moir, T. 2018. "Why is Implementation Science Important for Intervention Design and Evaluation within Educational Settings?" *Frontiers in Education* 3:61. <https://doi.org/10.3389/feduc.2018.00061>.
- Norwegian Data Inspectorate. 2022. "Data Protection Services" *Norwegian Data Inspectorate*. Accessed May 26, 2022. <https://www.nsd.no/en/data-protection-services>.
- Rusu, C., V. Rusu, S. Roncagliolo, and C. González. 2015. "Usability and User Experience: What Should We Care About?" *International Journal of Information Technologies and Systems Approach (IJITSA)* 8 (2): 1–12. <https://doi.org/10.4018/IJITSA.2015070101>.
- Saunders, B., J. Sim, T. Kingstone, S. Baker, J. Waterfield, B. Bartlam, H. Burroughs, and C. Jinks. 2018. "Saturation in Qualitative Research: Exploring Its Conceptualization and Operationalization." *Quality & Quantity* 52 (4): 1893–1907. <https://doi.org/10.1007/s11135-017-0574-8>.

- Schmidt, M., and R. Huang. 2022. "Defining Learning Experience Design: Voices from the Field of Learning Design & Technology." *TechTrends* 66 (2): 141–158. <https://doi.org/10.1007/s11528-021-00656-y>.
- Stewart, D. W., and P. Shamdasani. 2015. *Focus Groups: Theory and Practice*. Thousand Oaks, CA: Sage Publications.
- The Norwegian National Research Ethics Committees. 2022. "Guidelines for Research Ethics in the Social Sciences and the Humanities". *The Norwegian National Research Ethics Committees*, May 26. Accessed May 26, 2022. <https://www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/guidelines-for-research-ethics-in-the-social-sciences-humanities-law-and-theology/>.
- The International Organization for Standardization. 2017. "ISO/IEC 29341-28-10: 2017: Information Technology—UPnP Device Architecture—Part 28-10: Multiscreen Device Control Protocol—Application Management Service." *The International Organization for Standardization*, October. Accessed December 12, 2021. <https://www.iso.org/standard/69181.html>.
- The International Organization for Standardization. 2018. "ISO 9241-11: 2018: Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts". *The International Organization for Standardization*, March. Accessed December 12, 2021. <https://www.iso.org/standard/63500.html>.