# Thesis

# Certificate of Advanced Studies in Evaluation – University of Bern

## Development of Guidelines on "Evaluation in Times of Covid-19"

### Based on the analysis of recent publications and to date recommendations by the international evaluation community


Marina Häusermann M.A.

KEK-CDC, Universitätsstrasse 69, 8006 Zürich

16. February 1991, Zürich


Supervisor: Prof. Dr. Lars Balzer

Chair of the program: Dr. Stefanie Krapp

Interim study management: Marielle Schaer-Selby

Zürich, September 2020

## Abstract

The global Covid-19 pandemic impacted and transformed our everyday life and work. Projects and programs are adapted in response to changing needs and priorities. While evaluation is even more critical in times of crisis, field missions and face-to-face meetings are restricted to impossible. Evaluators are confronted with the questions on how to adapt their evaluation practice to these circumstances. Credible evidence is needed for decision-making and evaluators should try to generate it despite possible selection bias and other constraints.

The Covid-19 pandemic compelled an immediate response by the development evaluation community. As consultants at KEK-CDC are facing similar challenges and insecurities, a guideline on evaluations in times of Covid-19 was elaborated, based on a literature review and subsequent exchange with practitioners met at the Evaluation Hackathon and consultants at KEK-CDC.

There are many aspects to be considered. In summary, ethical principles are essential, including sensitivities to stakeholders and systems under pressure. Further, it is important to be flexible, do realistic planning and expect and facilitate change. Different ways of remote data collection have to be considered and finally, it is relevant to share results and experiences, to connect with other practitioners and to understand the current challenges as an opportunity of growth.

# Content

# 1 Background

## 1.1 Problem statement and objective of the thesis

> *"All evaluators must now become developmental evaluators, capable of adapting to complex dynamics systems, preparing for the unknown, for uncertainties, turbulence, lack of control, nonlinearities, and for emergence of the unexpected. This is the current context around the world in general and this is the world in which evaluation will exist for the foreseeable future."* (Patton 2020)

As Michael Patton wrote in late March, we are living in times of uncertainties, and as our environment, our societies and the frameworks we are working in are changing, we have to adapt our approaches and evaluation practice, as this is not a temporary state we are in, but – as Patton concludes, a condition that will remain. Due to the worldwide Covid-19 pandemic, field missions and face-to-face meetings are restricted to impossible. Therefore, some on-site evaluations cannot be carried out and evaluators have to look for different means to gather the necessary data. At the same time, projects and programmes are being adapted in response to changing needs, priorities, constraints as well as opportunities (Macfarlan 2020a). Choices being made now might shape our societies for years or decades to come and despite possible selection bias and other constraints, evaluators should take on the responsibility for generating credible evidence (Ofir 2020), while trying to avoid putting unnecessary pressure on a public system already heavily under pressure (ILO 2020; Vaessen und Raimondo 2020). Given the current situation, it is therefore important to consider how evaluation can be useful, what can be evaluated remotely, and which evaluation approaches and methods are particularly suitable for this purpose. The Covid-19 pandemic compelled an immediate response by the development evaluation community, leading to a huge amount of institutional publications, blogposts and individual reflections. Although or because we have no experience in how to cope with a crisis of this scale (J. R. Garcia 2020), it is unleashing huge potentials, including in the International Evaluation's Community (Weitzenegger 2020).

My colleagues at KEK-CDC, a consultancy firm advising public administrations and non-profit organizations in the Swiss and the international context, are facing similar challenges and insecurities. Therefore, the present thesis, written in the framework of the CAS program on evaluation, pursues the objective of elaborating guidelines on adaptions of current evaluation endeavors, which should serve my employer and colleagues at KEK-CDC as an aid and orientation in their everyday work and enable further reflections on evaluation activities in times of Covid-19. It is addressing evaluators and focuses on evaluations in the context of development cooperation, where reflections, exchange and knowledge-sharing have been most vivid. Due to travel restrictions, it was also the area of work raising most questions at KEK-CDC. Nevertheless, considerations about the availability of people and institutions in times of crisis, changed purposes or online data collection are of use to all evaluators.

## 1.2 Methodology

In order to elaborate the guidelines, the present thesis is reviewing and analyzing possibilities of adaption of evaluation to fit into the restricted framework of our new worldwide reality. Besides standard literature on evaluation in German (Balzer/Beywl 2018; Beywl 2007), publications on real-world (Bamberger/Mabry 2020) and developmental evaluation (Dozois u. a. 2011; Patton 2011) and evaluation in hard-to-reach contexts (ESS 2019) were screened. Furthermore, pre-Covid-19 literature on remote and online data collection as well as on Big Data was reviewed (Lune/Berg 2017; Petersson/Breul 2018; York/Bamberger 2020). In a second step, websites of multilateral organizations were systematically screened. Especially the World Bank, the Organization for Economic Co-operation and Development (OECD) as well as different UN agencies were elaborating short guidelines and publishing blog articles on the adaption of evaluations to Covid-19. These are considered fundamental sources for the elaboration of the guidelines through the present paper (IEO/UNDP 2020; ILO 2020; UNFPA Evaluation Office 2020). Additionally, webpages of national evaluation societies were looked at. The Working Group on Development Policy of the German Evaluation Society DeGEval (AK Epol-HuHi) for example created a new subgroup regarding the subject of remote evaluations[1] and realized an online event on evidence-based international cooperation (DEval/AK-Epol-HuHi 2020). Both the Australian Evaluation Society (AES)[2] and the American Evaluation Association (AEA)[3] published different blog articles on the subject and the European Evaluation Society made a compilation of useful articles and links (European Evaluation 2020).[4] Interestingly, the Swiss Evaluation Society SEVAL has not taken up the subject so far.[5]

Overall, a very intense practice of knowledge sharing can be observed, and the different articles, blogs and webinars are creating a fast-growing grid of ideas and recommendations. Within this thesis, the attention was put on publications by multilateral organizations and national evaluation societies and other actors in the field, such as the NGOs BetterEvaluation[6], CartONG[7] or the Blue Marble Evaluation Network.[8] Additionally, individual blogs and articles as well as podcasts[9] were reviewed to combine individual and organizational views. Nevertheless, both the timeframe and scope of this thesis are too small to take into consideration all newly appearing contributions. Furthermore, six months into the pandemic has been too short for a second wave of articles discussing in-depth lessons learnt and experiences made applying the proposed approaches and tools. It is therefore important to note that the outlined propositions and considerations about adapted evaluation processes due to Covid-19 should be seen as a snapshot, mirroring the current state of research.

The review and analysis of the mentioned publications are presented in chapter 2, structured alongside the three phases of an evaluation process according to Beywl (2007). At the end of every subchapter, a set of question is elaborated and proposed, which should help evaluators reflect on the most important aspects of adapted evaluation. Further, the aspect of online interviews and focus groups as well as the use of secondary and big data are deepened. Chapter 3 on the contrary is based on personal experiences and exchanges with other practitioners, complementing the insights from chapter 2. During the process of literature review, the very first Evaluation Hackathon took place

[1] Private Email from Susanne von Jan, AK Epol-HuHi, DeGEval, 30.08.2020.
[2] Australian Evaluation Society: https://www.aes.asn.au
[3] American Evaluation Society: www.eval.org, Blog: https://aea365.org/blog/
[4] European Evaluation Society: https://europeanevaluation.org/evaluation-in-times-of-covid-19/
[5] Private Email from Christian Rüefli, SEVAL, 30.08.2020.
[6] BetterEvaluation: https://www.betterevaluation.org/
[7] CartONG: https://www.cartong.org/
[8] Blue Marble Evaluation: https://bluemarbleeval.org/
[9] E.g. EvalCafé and the #EvalCrisis Podcast by DEVCO/ESS

(7.-13.7.2020). It was part of the online program, as the International program for development evaluation training (IPDET) could not take place physically in Bern. It united almost 600 people from 114 countries, organized within over 30 different teams. Motivated by Dr. Stefanie Krapp, head of the further education evaluation program at the University of Bern and head of IPDET, I submitted the subject of alternative data collection as a challenge to the evaluation hackathon (Annex 5)[10]. It was chosen as one of ten challenges, combined with a similar one on physical site observations in the context of Covid-19.[11] Personal interviews with partners at KEK-CDC and group work realized at the annual KEK-CDC retreat (20.-21.8.2020) add another layer of reflections. The considerations of chapter 3 were used to enrich and complement the key considerations of chapter 2, leading to the guidelines on adapted evaluation practices presented in chapter 4. The thesis ends with a personal reflection in chapter 5.

## 1.3 Terminology

In the blogs, articles and documents published during the last months, several terms appeared as designations for evaluations that could not be realized with a mission to the field. These include the terms *decentralized evaluation, off-site, remote or adapted evaluation*. Nevertheless, a distinction is sometimes made between *semi-remote/hybrid* and *remote evaluations*. The semi-remote/hybrid evaluation includes online-interviews carried out by the international consultants and face-to-face (f2f) interviews and interactions realized by the national consultant or by the international consultant at a later point (two-phased data collection) (cf. ILO 2020, 3). On other occasions, the term *mixed-modality* has been used for semi-remote / hybrid data collection (cf. FAO 2020, 6). The *remote evaluation* on the other hand includes only online data collection and is conducted if the mobility restrictions or security issues prevent also the national consultant to carry out f2f interviews and visits to the field (cf. ILO 2020, 3). Within this paper, the terms semi-remote, hybrid and adapted evaluation are used interchangeably.

A glossary of English and German evaluation terms based on Beywl and Niestroj (2009) can be found in Annex 3.

[10] Project page of the Team Phoenix: https://evalhack.org/project/46
[11] See all the 2020 evaluation hackathon challenges on https://ipdet.org/online-program-2020/ipdet-evaluation-hackathon/challenges/.

# 2 Adapting evaluation in times of Covid-19

*Healthy breathing is about balance.*

*Evaluation is the oxygen that powers decision making. Too little and we are likely to make poor decisions. And when faced with big challenges, we need more than usual.*

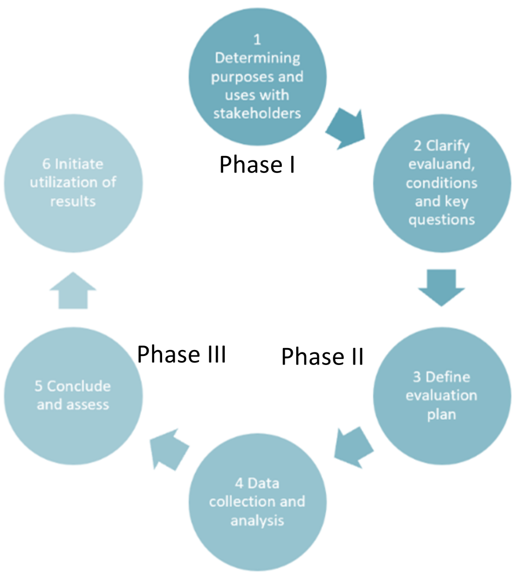
*Too much evaluation without action leads to hyperventilation. Analysis paralysis.*

*As an evaluator, it is your responsibility to keep the breathing steady.*

(Lysy 2020)

Evaluation is the systematic use of scientific research methods, techniques and theoretical knowledge to assess an object with regard to decision and operational situations (Ritz 2003, 29). The analogy to healthy breathing by evaluator and cartoonist Chris Lysy is reminding us of the fact, that despite of all difficulties, evaluation work has to be continued, taking into account both limitations and opportunities of growth and learning (cf. Parker 2020). Nontraditional evaluation approaches can support decision making and reflection in complex circumstances (Josephson et al. 2020). Patton stressed the importance of developmental evaluations (Patton 2020), an approach which emerged to support real-time learning in complex and emergent situations and facilitate change (Dozois u. a. 2011). In times of uncertainties, points of orientation are key and as Jade Maloney from AES pointed out, we can take comfort in the fact that although how we evaluate might be changing, what is at the core of evaluation is not (2020). At the core is a process, which according to Beywl (2007) can be divided into three phases and six steps (Figure 1).

*Figure 1: Evaluation steps according to (Beywl 2007).*




Phase I (identification of the evaluand) starts from taking on the evaluation mandate on determining purposes and intended uses of the evaluation as well as its scope, budget, and timeline. Further, this subchapter includes reflections about the involvement of stakeholders, the clarification of the evaluand and the prioritization of key evaluation questions to be answered, leading to the design of the evaluation plan. This second phase (information acquisition) is based on the design of the data collection, including both the following data collection and analysis. The subchapter further deepens on online interview and focus group on one hand, and the use of secondary and big data on the other. The chapter will be closed focusing on Phase III (communication of findings) on the conclusions and utilization of results (Beywl 2007, 7). Some initial thoughts will be given on feasibility and ethics.

## 2.1 Considering feasibility and ethics

Although the present thesis is addressing evaluators, important steps in the planning realized by commissioning bodies are mentioned if they change due to Covid-19. To obtain useful results from an evaluation, it is key to specify what shall be evaluated – the evaluand – before calling for tenders. In the further process, the evaluand is defined or redefined jointly by evaluators and stakeholders, which will be discussed in chapter 2.2. Even though it is uncertain, when the situation will 'normalize' again, commissioning bodies might decide to cancel or postpone evaluations.

Before engaging in an evaluation process, the **feasibility** of the evaluation should therefore be assessed. The FAO suggests, that if risks in terms of health are high, whereas risks in terms of quality are manageable and the evaluation has priority, it should be conducted in a two-phased modality (2020). Ideally, both the Terms of Reference (ToR) and the offer should include an analysis of potential Covid-19 related risks to the evaluation and a display of considered mitigation strategies (DEval/AK-Epol-HuHi 2020). The evaluation team is meanwhile confronted with the question if it still has the necessary skills, if the design of the evaluation is changed and which reduced or additional costs this would imply. To capitalize resources and allow for semi-remote data collection, the evaluation usually relies on national consultants. The use of in-country expertise is already a key aspect of the business-as-usual scenario but is gaining importance because of the mobility restrictions (Raimondo, Vaessen, und Branco 2020). Responsibilities are divided differently, as more tasks are carried out by the national consultants, which due to the increasing demand might be less available (ILO 2020, 6). At the same time, these adaptions offer opportunities for strengthening and revaluating local evaluation expertise as well as of a more collaborative way of knowledge production (Open Access Repositories 2020; Ugarte 2020). Aspects regarding the feasibility of semi-remote data collection will be deepened in chapter 2.3.

From an **ethical** point of view, considerations of the risk-reward ratio of evaluation activities are gaining importance. The quality standards of SEVAL mention ethics in a general manner (SEVAL 2017). Most recent publications specifically highlight the application of the principle of "do no harm", ensuring the well-being and safety of national and international consultants as well as other involved stakeholders during the whole process of the evaluation (FAO 2020; IEO/UNDP 2020, 6, 9; Raab 2020). Another guiding principle is "leave no one behind", keeping in mind vulnerabilities and inequalities. Evaluation should be human-rights-based, equity-focused and gender-responsive (UNFPA Evaluation Office 2020; Weitzenegger 2020). Therefore, it is important to think about different possibilities to ensure that all groups of stakeholders are able to have a voice during the adapted evaluation process. The following table summarizes the key considerations on the two discussed topics, proposing questions evaluators should reflect on.

*Table 1: Key considerations on feasibility and ethics*

| | |
|---|---|
| **Feasibility** | • In as far has the evaluand changed? Can conditions be met for the evaluation to be useful and used?<br>• Are the current project/ program objectives and outcomes coherent enough to allow for evaluation?<br>• Can the evaluation be conducted without undue stress to stakeholders responding to Covid-19?<br>• What are the additional or changed roles, processes or approaches, that you will need to address? Are you comfortable carrying out or facilitating these |

| | |
|---|---|
| | changes? |
| | • Does your team still have the necessary skills if the evaluation design has to be changed? If not, how can you bring in additional external expertise or strengthen the capacity of your team?<br>• Are qualified local consultants available to support your team? Do local Covid-19 restrictions allow for in-country mobility and face-to-face-meetings?<br>• Did you consider taking on a mentoring role for the national consultant? |
| **Ethics**  | • In what ways could your evaluation put team members, local consultants, and stakeholder at (health) risk and how can these risks be mitigated?<br>• How can you ensure that all your groups of stakeholders are able to have a voice during the evaluation process? How will any adaptation affect your most marginalized stakeholders (e.g. access to technology)?<br>• How can you assure that the adapted design of the evaluation is considering questions of equity and gender-responsiveness? |

*Source: Own compilation of questions based on cited literature*

If the commissioning body and the chosen evaluating team agree that an evaluation is feasible, the process of the evaluation can be started with phase I, the identification of the evaluand.

## 2.2 Phase I: Rethinking priorities and adapting approaches

To ensure the results will meet the information needs of stakeholder and appropriately mirror their interests, **stakeholder identification and engagement** are key to an evaluation. Consideration must be given to the question, how the crisis has affected key stakeholders and intended users of the evaluation, as availability and needs may have changed (Macfarlan 2020a). Due to mobility restrictions and shifting institutional priorities, unequal skills and access to digital tools, some of the key stakeholders might not be available for interviews, leading to a selection bias. They might also have additional duties or roles, hindering their participation in and commitment to the evaluation. As in every evaluation process, the main **purposes** and intended uses of an evaluation must be reviewed together with the stakeholders (Beywl 2007). In the current situation, it is even more important to clarify, whether the prioritization of the initial purposes of the evaluation is still the same and if there are any new uses for the evaluation that have recently emerged. Both would lead to changes in the approach, the design and the deliverables of the evaluation. The evaluation could include both summative and formative elements, e.g. how the project or program performed in terms of past achievements and if they are at risk (summative) and how the project / program has been adapted or could be repurposed to the crisis (formative) (ILO 2020, 6).

As mentioned, organizations around the world were forced to adapt their project and program activities in response to changing needs, priorities, constraints and opportunities and evaluation needs to match this (Macfarlan 2020a). As public resources are diverted to the Covid-19 crisis, some of the planned interventions may not be implemented as designed. Evaluation plans have to be rethought in order to respond to upcoming practical considerations and reprioritize the most urgent evaluation questions (IEO/UNDP 2020, 6).

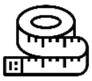
**Defining the evaluand**, program trees, logical frameworks and theories of change can be useful tools to develop a common understanding of the logic of intervention of the project or program (Beywl 2007, 22 ff.). Patton urges evaluators to actively support the ongoing changes and adaptations and to support stakeholder to deal with what is unfolding. “Expect change. Facilitate change. Document changes and their implications. That’s your job in a crisis […] And if you don’t see program adaptation, consider pushing for it by presenting options and introducing scenario thinking at a program level” (Patton 2020). Once the purposes and the evaluand are set, **key questions** are formulated. As in all evaluation endeavors, it is important that the evaluation questions fit the purpose of the evaluation and the information needs of the involved stakeholders.

After having set the framework and key questions, values to be used in the evaluation in terms of **criteria** (aspects of performance) and **standards** (levels of performance) need to be clarified. The evaluation of international programs is guided by the evaluation criteria formulated by the OECD Development Assistance Committee (OECD/DAC), elaborated in 1991 with the most recent revision dating from 2018 (OECD/DAC 2019). Therefore, the evaluation will probably still use the six defined criteria, namely Relevance, Coherence, Effectiveness, Efficiency, Impact and Sustainability. Nevertheless, the standards might need to be adapted, as new domains of performance have become more important, it is not reasonable anymore to expect the same standards of performance due to changed circumstances and because it might have been redefined how success looks like (Macfarlan 2020b).

Derived from the literature review, key considerations and questions which evaluators should ask themselves are summarized in the following table. It is structured alongside the bold keywords highlighted in the text. The selection of keywords is based on Beywl’s publication on the planning of evaluations (2007).

*Table 2: Key considerations rethinking priorities and adapting approaches*

| | |
|---|---|
| **Stakeholder Engagement** | • How has the Covid-19 crisis affected the key stakeholders of your evaluation?<br>• Have their roles changed? Do they have additional duties that will take away their focus from engaging with the evaluation? |
| **Purposes** | • Is the prioritization of the initial purposes of the evaluation still the same?<br>• Are there any new uses for the evaluation that have emerged in the current circumstances?<br>• Is there a potential change in deliverables?<br>• Are stakeholders likely to listen and able to act on the findings and recommendations of the evaluation? |
| **Definition of the evaluand** | • In as far has the program or project been changed due to Covid-19?<br>• In as far has the theory of change/ Logical framework therefore to be revised?<br>• How can you support stakeholders in the process of adaptation (e.g. scenario thinking) and document change and its consequences? |

| | |
|---|---|
| **Evaluation questions**  | • Have the evaluation questions been adapted to newly formulated purposes and information needs?<br>• How can the aspect of equity be reflected in the key questions? |
| **Criteria**  | • How might the standards and criteria used in the evaluation need to change?<br>• Has it been redefined, how success looks like? |

*Source: Own compilation of questions based on cited literature*

## 2.3 Phase II: Design of evaluation plan and (semi)remote data collection

Following Beywls layout, the second phase of the evaluation process includes the evaluation design as well as data collection and analysis. It is challenging to choose the appropriate **evaluation design**, as there is not a single perfect solution, not even for “a particular program, at a particular time, with a particular budget” (Cronbach/Shapiro 1982, 231). It depends on the questions of the evaluation as well as on the feasibility, data security, access to stakeholders and to the field as well as the already mentioned budget options (cf. Balzer und Beywl 2018, 92 ff.), even more so as currently some of these points cannot be taken for granted and have to be closely analyzed. Usually, field missions are at the very core of project and programmatic evaluations. While the work starts with document analysis, a review of available information, it is by interacting with policymakers, implementing agency and beneficiaries, that the evaluator gets a broader understanding of the reality affecting the design and implementation of projects (Vandercasteelen 2020). Travel restrictions and social distancing measures ask for an adaptation of the data collection approach. Ideally, the data collection is realized in a two-phased way, meaning that online/remote data collection is combined with f2f interviews and interactions by either a local consultant or the international consultant at a later stage. To allow for triangulation, different methods of data collection such as surveys, desk-review of secondary information, (online) interviews etc. shall be combined (cf. ILO 2020, 8; UNFPA Evaluation Office 2020). This matters for all evaluations but is even more important regarding the selection bias and limitations of online data collection. Patton as well as other authors stressed the fact, that information is needed quickly to help stakeholders make decisions and that due to the current situation, sometimes “good enough” data has to be sufficient (Josephson et al. 2020; Patton 2020). Nevertheless, similar conditions for remote data collection could also arise for other reasons: Be it due to a critical security situation in the destination country or budgetary or ecological considerations. Based on the evaluation questions, the evaluation matrix is created, identifying the type of stakeholders that would be best positioned to provide inputs on the specific questions. While setting the methodology for each of them, it should be reflected, whether online methodologies are justified and appropriate to provide sufficient data. Adaptions have to be thought through, because to move the exact same methods and instruments online might not always be the best option. This could lead to a decision not to address particular questions or considering new ones (ILO 2020, 7). As mentioned in Phase I, considerations have to be made if informants can be accessed and what is needed to maximize participation and inclusiveness through remote means, and time and resources allocated (UN Women 2020).

CartONG[12] as well as Fabio Bezerra from Unicef Mozambique elaborated overviews on remote and hybrid **data collection** (Bezerra 2020; C. Garcia 2020). They recommend limiting f2f data collection to minimize health risks for evaluators and informants. Further they highlight the importance to focus on essential and critical data needed for the evaluation and to make use of already existing secondary data to limit unnecessary data collection (chapter 2.3.2.) For this reason, evaluators as well as organizations and institutions are encouraged to share their data (C. Garcia 2020). However, this raises questions about data privacy and security.

In the case of adapted evaluations, realistic **planning is key**. Contrary to wide-spread assumptions, the set-up, preparation and documentation of online data collection such as interviews might make up the time gains from not travelling. In order to keep the possibility of hybrid evaluation, it is important to negotiate longer evaluation timeframes to allow phased data collection, e.g. a first phase of remote, a second phase of f2f interviews by the international or the national consultant (cf. ILO 2020, 7–8). The IFAD Independent Office of Evaluation (IOE) suggests complementing visits to the field by national consultants through validation missions by the international consultants. On the other hand, the time stretching of ongoing evaluation projects might also be a challenge. Furthermore, the setting of an international consultant who is remotely leading an evaluation, and a national consultant realizing a great part of the data collection, generates a huge need for communication and coordination. Obviously, this need exists, even if both consultants are at the same geographical place. The difference is that if they are not, specific time slots have to be planned for exchanges and (daily) debriefings. Due to the uncertainty of global developments, it is further recommended to do a regular reassessment of possible adaptions (DEval/AK-Epol-HuHi 2020). Planning also includes considerations about additional costs for translations, tools for remote data collection and the documentation of learnings from the adapted procedure. The key considerations regarding data collection are summarized in the following table. Questions regarding interviews, focus groups and the use of secondary and big data will be discussed in the following two subchapters.

*Table 3: Key considerations on the design of evaluation plan and (semi)remote data collection*

| | |
|---|---|
| **Evaluation design**<br> | • Do you have the capacity and resources to collect and analyze the data needed to answer the evaluation questions? If not, can you capitalize on existing data or acquire the necessary skills?<br>• Are you likely to have well-substantiated evaluation findings? What practical and ethical constraints are weighing on your capacity to collect information?<br>• Which stakeholder can provide information to which evaluation question?<br>• Can you adapt your methods to comply with physical distancing / mobility restrictions?<br>• Are online methodologies appropriate to gather valuable data for the particular evaluation questions?<br>• Did you consider elements of approaches such as developmental evaluation? |

[12] CartONG is a French NGO with the objective to provide mapping and information management services to humanitarian and development organizations, and to promote the use of Geographic Information System.

| | |
|---|---|
| **Planning**<br> | • Has the evaluation budget been adapted? What additional/reduced costs will you have?<br>• Have you budgeted for capturing learnings about this process?<br>• Did you plan for coordination, documentation of interviews and (daily) debriefings with your national consultant?<br>• Is there flexibility to adapt to new uncertainties and changes in the overall framework?<br>• Is there a possibility for a two-phased approach to data collection?<br>• Did you consider possible additional work and expenditure for translations? |
| **Data collection**<br> | • Do reliable contact lists exist, guaranteeing access to stakeholders/interviewees?<br>• If you move your activities from f2f to online, what conditions or support will be needed to make participation and engagement possible?<br>• Is internet connectivity granted or what alternatives of remote data collection should be considered?<br>• Can you include more desk reviews and make use of existing secondary data or even big data (e.g. GIS data)?<br>• Do you have the necessary skills to work with existing and accessible big data sets? If now, are you considering to acquire them? |

*Source: Own compilation of questions based on cited literature*

### 2.3.1 The new normal: Online interviews and focus groups

There are myriad ways to gather qualitative information online (Jowett 2020), out of which online interviews and focus group discussions[13] have seen a huge increase during the Covid-19 pandemic. However, they have already been used by researchers and evaluators starting in the late 1990s, while in-person f2f interviewing remained the dominant and most used interview technique and believed 'gold standard'. While interviews could be held over the phone, online focus groups were initially based on e-mail or web-based messaging. The last decade showed an increased use of existing technologies accompanied by a rapid development of new ones, offering alternative modes such as Voice over the Internet Protocol (VOIP)[14] systems (cf. Hinchcliffe/Gavin 2009). Technology keeps pushing the boundaries of online collaboration, enabling researchers and evaluators overcome time and financial constraints, geographical dispersion, physical mobility boundaries and reducing staff requirements (Barratt/Maddox 2016, 1; Hagan 2018; cf. Janghorban et al. 2014). The current situation due to the Covid-19 pandemic further strengthened the digital skills of people worldwide who were forced to work from home and transfer tasks to the online world.

The **online interview** modes are differentiated by their (a)synchrony in time and space (cf. Lune/Berg 2017, 80; cf. Opdenakker 2006). Asynchronous interview techniques like e-mail reduce spontaneity and might therefore influence the answers of the respondent, because he or she can prepare them. Synchronal techniques enable an immediate experience of the interview situation and can be divided into the ones using only written, oral or also visual elements (Ahsan 2020). Text and audio messaging can be used either synchronously or asynchronously (cf. figure 7 in Annex 4). The more dimensions

[13] The focus group is a well-used interview style suitable for small groups, formed by the evaluator/researcher addressing a particular topic to generate insights through group discussion and interaction (cf. Lune/Berg 2017, 94 ff.).
[14] Many of the reviewed studies of older date focus on Skype as main tool for realizing online-interviews. In this paper, other videoconferencing tools such as Zoom, Webex or Jitsi are treated equally.

are added, the richer the experience gets, as it might include the pace of the answers, the tone of the voice, facial expressions and body language as well as an impression of the surrounding area the interviewee is in. On one hand, online techniques allow for interviews with different stakeholders in different parts of a country or region, which – due to the longer time frame of the evaluation – might have more time and be better prepared for the interview (DEval/AK-Epol-HuHi 2020). Nevertheless, it depends on the topic and the circumstances, which approach or combination may be most adequate and the assumed superiority of the f2f interview is not demonstrably absolute (cf. Krouwel et al. 2019). Some researchers argue that since phone interviews, online focus groups and online surveys provide a kind of anonymity, they are effective to obtain “hard-to-locate individuals or when asking highly sensitive questions (Barratt/Maddox 2016, 3; Champion 2006).

Concerns of authenticity in interactions always remain, but according to Jessica Sullivan we should not “question the truth in these interactions as we are in search of the meanings that the participants find and assign” (Sullivan 2012, 59). On the other hand, online interviewing makes it harder to build rapport and trust with the interviewee and lacks the possibility of unobtrusive observation (Raimondo et al. 2020). Even though interviews might be to the point and gather the necessary information, there is no space for the usual informal chatting, revealing valuable insights, as other evaluators confirmed (DEval/AK-Epol-HuHi 2020). Live streams of interview settings and surroundings might give a broad impression but are not able to fill the gap. Berg et al. (2017) outlined, that it is easier to work with and online interview people, you have already seen in person. Even though it might be easier to find a convenient time slot, one has to keep in mind, that people might be distracted due to lockdown situations or care and work duties (UN Women 2020). Regarding the setup of the online interview, they as well as other authors stressed that it is even more important for online interviews, that the evaluator is following a formal or semi-structured interview guideline, which is kept short, asking straightforward questions. This is also true for online focus groups, where clear ground rules should be established, and it is recommended to use slides to display topics and questions (cf. figure 5 in Annex 4).

As well as with interviews, there are synchronous and asynchronous ways to conduct **online focus groups** and they include modern communication channels such as VOIP or instant messengers such as WhatsApp (cf. Chen/Neo 2019). Chen and Neo (2019) outlined that WhatsApp focus groups for example have the advantages that accessibility levels of smartphones and WhatsApp are very high and that especially younger people, often referred to as ‘digital natives’, even prefer communicating digitally. On the downside of online focus groups is the loss of the role and authority of the moderator as well as the inability to sense the atmosphere properly, which makes it hard to effectively use group dynamics as an integral part of the overall process. It needs a strong facilitator, taking into account the questions of interest as well as the totality of participants. Therefore, the groups should be kept smaller than in person (Forrestal et al. 2015). Another point to be kept in mind is the language issue. If the focus group has to be held in a foreign language, the responsibility of facilitation lies with the local consultant. There is still the option to have the international consultant within the focus group discussion l, getting a simultaneous translation, in order to be able to follow and partly guide the direction of the discussion.[15] While it seems obvious, that the collection of primary data has to be adapted to the situation, we also have to consider an adapted use of secondary or even big data.

*Table 4: Key considerations on online interviews and focus groups*

| | |
|---|---|
| **Online interviews and focus groups**  | • Which factors may limit participation (connectivity, skills, language, availability)?<br>• Which kind of online interviews/focus groups are most appropriate for tackling the present topic / addressing the present stakeholders?<br>• Is there a possibility to do a live stream of the interview settings and surroundings to get a visual impression and relate to the interviewee?<br>• Do you have the resources to do more focus groups with less people?<br>• Have you set the ground rules for discussion to facilitate the focus group?<br>• Do you have a private and comfortable space to engage online? |

*Source: Own compilation of questions based on cited literature*

### 2.3.2 New frontiers: The use of secondary and big data

In regard to the challenges we are facing, it is important to consider not only, how interviews can be held online, but also how we can use modern technology to collect or make use of data, which has already been gathered by other means. In any case, evaluators are increasingly acting as reviewers and synthesizers of existing knowledge (Raimondo et al. 2020). Whereas the use of secondary data through desk research, project records and official statistics is widespread, the rise of new technologies has been neglected (York/Bamberger 2020, 25) and there are only a few documented examples, where Big Data[16] has been rewardingly used in evaluation.

Already in 2017, Gustav Petersson and Jonathan Breul stated that "increasing complexity is pushing evaluators to move from simple linear models to a more realistic way of evaluating development processes and their outcomes" and discuss how "societal and technological changes provide unprecedented opportunities and challenges for evaluators" (2018, vii). Technology should accordingly be seen as an enabler rather than a goal in itself, offering the following advantages: It can collect and analyze data rapidly, offers the possibility to collect a variety of data points, saves costs and might increase capacities and collaboration (Petersson/Breul 2018; York/Bamberger 2020). On the downside, a specific set of skills is needed and there is a risk of selection bias, e.g. due to unequal access to technology (Bezerra 2020). Moreover, it has to be kept in mind, that the data was collected for a "transactional, documentation or tracking purpose" (cf. York/Bamberger 2020, 16) and that it therefore differs from evaluation data gathered for a research purpose, e.g. investigating and measuring previously defined key concepts and variables. Big data as related to evaluation can be divided into three categories: human-generated (centered), administrative (transactional), and geospatial data, the latter becoming increasingly present in evaluation, as it can provide valuable insights for the assessment of project effectiveness, as for example regarding reforestation (DEval/AK-Epol-HuHi 2020; Vandercasteelen 2020). While big data already exists and might serve as secondary data to be included and complement the data collected specifically for the evaluation, new

[15] This alternative was used by a colleague at KEK-CDC in August 2020.
[16] Big data consists of "(call) logs, mobile-banking transactions, and online user-generated content, such as blog posts, social media, sensor apps, online searches, and satellite images, including all the data that one holds in e-mails, weblogs, intranets, extranets, etc." (cf. Petersson/Breul 2018, 73).

technologies can also be used to specifically collect primary data needed, as outlined by Bezerra (2020). Jeff Chelsky and Lauren Kelly suggest that technology tools used for digital data collection during Covid-19 should be designed in line with data collection requirements and combined to allow for triangulation of data. As for all online data collection, exclusionary factors have to be considered as well as the fact that tech-enabled tools might be excellent at verifying assets, but are not adapted to measure quality or implementation fidelity (cf. Figure 6 in Annex 4). Therefore, the human element in data collection techniques should not be underestimated (cf. ILO 2020, 2). It is hence valuable to glimpse at the possibilities offered by big data and technology, while keeping in mind its limitations. According to York and Bamberger (2020), all development partners should work together to build bridges between data scientists and evaluators. Maybe the challenges set by Covid-19 lead to a faster adoption of tools and techniques of data science.

*Table 5: Key considerations on the use of secondary and big data*

| | |
|---|---|
| **Secondary and big data**<br> | • Can you include more desk reviews and make use of existing secondary data?<br>• Do you have the necessary skills to work with existing and accessible big data sets (e.g. GIS data)? If not, is it worth investing in these skills?<br>• Do these sources of (big) data help you answering particular evaluation questions?<br>• Can secondary and big data be used ethically? |

*Source: Own compilation of questions based on cited literature*

## 2.4 Phase III: Conclusions and utilization of results

Although there are guidelines and points of orientation regarding the adapted planning and data collection for evaluations, publications regarding conclusions, recommendations, and the utilization of results are scarce. Various authors suggest that the findings and recommendations of an evaluation are validated and discussed with stakeholders in virtual meetings (IFAD 2020). Due to the current uncertainty about the evolution of events, it is recommended to increase feedback and reflection loops between evaluators and clients to be able to adapt rapidly to possible changes in the overall setting of the evaluation. Furthermore, there is the question of fairness and usefulness of recommendations in this volatile context, as some might not be taken up due to the current circumstances. It is recommended to use well-structured briefs, infographics and shorter reports, as they are user-friendly and provide concise and practical information in times of content-overload (IEO/UNDP 2020). The sharing of both empathy as well as experiences and results across the international evaluation community is stressed (Ofir 2020). They further highlight the importance to stay connected with other evaluation offices, networks and associations, as collectively we can support each other to evolving evaluation approaches, methods, technologies and tools that are needed to continue our work during crises (DEval/AK-Epol-HuHi 2020; IEO/UNDP 2020). As the pandemic continues and experiences with adapted evaluations grow, the list of publications and possible key considerations will expand as well.

*Table 6: Key considerations regarding conclusions and the utilization of results.*

| | |
|---|---|
| **Conclusions and utilization of results**<br> | • Are the recommendations useful and can they be taken up according to the current circumstances?<br>• How can findings be presented concisely and user-friendly in times of content-overload?<br>• How can experiences (lessons learnt) be shared with the (international) evaluation community?<br>• How can we connect with other practitioners, learn and support each other?<br>• How can we advocate for evaluation in order to maintain and strengthen its profile? |

*Source: Own compilation of questions based on cited literature*

# 3 From literature to practice

As explained in the introductory part, the gathered findings and key considerations based on the reviewed literature are subsequently mirrored and complemented by the experiences of the evaluation hackathon as well as conversations and group work with colleagues at KEK-CDC.

## 3.1 Insights from the first evaluation hackathon

In a situation of crisis as the one we are facing, it is important to reflect on changed priorities and circumstances. As Mafcarlaan (2020a) stated, evaluation has to match this need, and so does the evaluation community. The evaluation hackathon represented an innovative format to respond to these needs. Participants were challenged to collaborate with people they did not know and who were spread over various continents and time zones, having different backgrounds and mother tongues, mirroring today's new ways and challenges of online collaboration. As in evaluation endeavors, understandings and objectives had to be discussed, aligned and defined first. Team Phoenix, I belonged to, was initially focusing on specific limitations of remote data collection, such as accessibility and selection bias. During the week and the daily discussions, the attention moved from specific methods and techniques applied to elaborate a platform to support evaluators and other stakeholders in the process of deciding which were the appropriate tools for which topic and evaluation context. Interestingly, this direction of solution was mirrored by other working on other challenges. In the focus was the idea of online platforms, connecting local to international consultants and provider of secondary data as well as on an exchange opportunities about lessons learnt. This aligns with the vivid knowledge sharing activities reflected in the blog articles and guidelines cited in chapter 2. As we will see in the following subchapter, it also mirrors the need for orientation and guidance brought up at KEK-CDC.

## 3.2 KEK-CDC Retreat: Thinking about challenges and opportunities

Evaluations in times of Covid-19 pose challenges, but also offer opportunities. The following have been gathered and discussed in the KEK-CDC retreat; The documentation can be found in Annex 6. Regarding feasibility colleagues confirmed, that it makes a huge difference, whether talking about a totally remote evaluation or a hybrid one, where local consultants are still able to go to the field for interviews or focus groups and workshops with small groups of people to experience and understand the setting and the group dynamics. They further confirmed Lune and Berg's comment, that it is easier to remotely work with and interview people you have already met in person. The same is valid for the context; Colleagues stated, they could partly close the gap of not experiencing the context by previous experiences of the same setting. On a more general note, questions treated in chapter 2.2. came up: Are stakeholders still available and interested in the evaluation? Is the project still working? How shall we evaluate the recent changes in the project and how do we handle the adaption of the evaluand? These questions are key. However, there is no straight answer to them, as they have to be discussed and clarified with the commissioning bodies and stakeholders. However, as Patton outlined, it is most important to expect and facilitate change and support stakeholders in these processes (Patton 2020).

Additionally, as independent evaluator Marie-Carin von Gumppenberg as well as others mentioned, planning was stressed as an important aspect. The mobility restrictions often do not lead to the assumed time gains. Colleagues stated that the data collection process takes much longer than on-site, where one has only a very limited time budget. Reflections on interviews and de-briefings have to be scheduled, translated scripts by the local consultant have to be revised together, and follow-up questions need to be formulated. Based on previous experiences, they added that remote data collection is easier regarding hard facts e.g. quantifiable data. This has also been a point made by different authors regarding the use of technology in remote data collection. Therefore, it remains difficult to get verified assessments about quality and implementation fidelity for example. In general, it was mentioned that the verification of data was difficult. Regarding online data collection, they further confirmed the difficulty to relate to interviewees and to notice nuances and underlying significances. They outlined the difficulty of creative group processes and network formation and of facilitating hybrid formats, where some people are f2f, others online. The complementation with secondary and big data is seen as an asset, but it is admitted, that previous experiences and skills are lacking. There were also some risks mentioned: Alongside the already discussed selection bias due to unequal access to technology for example, it was highlighted, that the dependency on clients to share contacts of interviewees could lead to limited flexibility.

Regarding opportunities, colleagues stated it had been uncommon until March 2020 to do hybrid evaluations and there was great skepticism towards online interviewing on part of the commissioning bodies. As mentioned above, “do no harm” is the most important principle now, and everyone is flexible and open to adaptions, hoping to make things work in one way or the other. Most of the opportunities were seen in the organization of work, the work-life-balance and the possible parallelism of different mandates. To have local evaluation expertise built-up and strengthened is also seen as an asset. At the same time, it was mentioned, that sometimes it might be easier for the international consultant as an outsider to the context and setting than the local one belonging to it, to voice criticism. In summary, colleagues at KEK-CDC shared the points put forward by the reviewed publications, while stressing the changed modes of organization of one's own work. “Expect change. Facilitate change. Document changes and their implications. That's your job in a crisis,” wrote Patton in March (2020). Following this call, the elaborated guidelines are an attempt to document and facilitate change in the evaluation practice due to the Covid-19 pandemic and its implications.

# 4 Guidelines on “Evaluation in Times of Covid-19”

Covid-19 led to uncertainties and insecurities that permeate our everyday life and work and as evaluators, we will have “to navigate multiple trade-offs, put ethics front and center, be willing to get out of our methodological comfort zones, and be ready to try and fail” (Raimondo et al. 2020). The questions in tables 1-6 serve as inputs to reflect on and think through the planning and realization of an evaluation. The following guideline provides recommendations on how to act and will be continuously adapted and revised, based on new experiences and learnings. It will grow with us. The challenges ahead should therefore not be seen as a discouragement but as an opportunity to grow. As Kirsten Wienberg said, we are now learning for the “new-normal world” (DEval/AK-Epol-HuHi 2020).

| | | |
|---|---|---|
| General | **Feasibility** | • Consider if conditions can be met for the evaluation to be useful and used.<br>• Include mitigation strategies in your offer.<br>• Reflect on the available skills of team members and if you feel comfortable enacting the expected changes. |
| | **Ethics** | • Do not expose yourself or other stakeholders to health and other risks.<br>• Be sensitive to increasing pressure on systems.<br>• Limit new or enhanced biases.<br>• Ensure that all groups of stakeholders have a voice in the evaluation process. Pay attention to make your evaluation equity-focused and gender-responsive.<br>• Take responsibility for creating credible and well-substantiated evidence. |
| Phase I | **Stakeholder engagement** | • Consider changed roles and availabilities of stakeholders.<br>• Be sensitive to stakeholders under economic and psychological strain.<br>• Be inclusive but focus on those who can provide substantive information and feedback. |
| | **Purposes** | • Reprioritize purposes together with stakeholders.<br>• Clarify if there is a change in deliverables.<br>• Reflect if stakeholders are likely to listen and able to act on the findings and recommendations of the evaluation. |
| | **Definition of the evaluand** | • Expect change and facilitate coping with uncertain situations as well as new vulnerabilities, tensions and lack of control.<br>• Present options and introduce scenario thinking to facilitate adaptions. |
| | **Evaluation questions** | • Adapt evaluation questions to newly formulated purposes and information needs.<br>• How can the aspect of equity be reflected in the key questions? |
| | **Criteria** | • Discuss standards of performance.<br>• Reflect with stakeholders on how success is defined. |

| | | |
|---|---|---|
| Phase II | **Evaluation design** | • Consider what evaluation scope can bring the most value.<br>• Look into nontraditional approaches such as developmental evaluation.<br>• Reflect on practical and ethical constrains weighing on your capacity to collect information.<br>• Adapt your methods to comply with physical distancing and mobility restrictions.<br>• Be purposeful in choosing online methods. |
| | **Planning** | • Do realistic scheduling and include time for briefings, exchanges as well as work breaks.<br>• Regularly reassess possible adaptions. Plan for regularly exchanges with commissioning bodies to reflect on adaptations and learnings.<br>• Assess the possibility of a two-phased data collection.<br>• Plan for coordination, documentation of interviews and (daily) debriefings with your national consultant. Consider taking a mentoring role for national evaluators if necessary.<br>• Consider additional resources for translations. |
| | **Data collection** | • Increase the use of new and digital technologies.<br>• Limit primary data collection to a minimum and focus on areas with data gaps.<br>• Choose the kind of interview that most suits your purpose / your stakeholders.<br>• Create conditions to allow for online participation.<br>• Consider doing more focus groups with less participants.<br>• Keep the process brief and focused on key issues, asking straightforward questions.<br>• Unlock the potential of desk reviews and analysis and capitalize on existing sources of data.<br>• Explore the possibility of accessing big data. Look for the necessary skills or start developing them. |
| Phase III | **Conclusions and utilizations of results** | • Document changes and their implications.<br>• Formulate useful recommendations which can be taken up by stakeholders despite uncertainties.<br>• Present findings concisely and user-friendly.<br>• Connect with other practitioners and share lessons learnt. |
| Continuously | **Learning** | • Schedule/budget for capturing learnings about the adapted evaluation process.<br>• Be courageous, try new approaches and acquire new skills.<br>• Facilitate intentional reflection on remote working processes.<br>• Advocate for evaluation and share experiences across the international evaluation community. |

# 5 Reflection

While my theoretical knowledge was growing due to the CAS-classes during spring 2020, Covid-19 seemed to slow down my application of newly gained skills. At the same time, Covid-19 changed the setting and framework of evaluation. New challenges arose and as the discussions and multitude of blog articles showed, there was a need for new practices and experiences – which made me curious. For the first time, I felt there was an international evaluation community, there were so many diverse and interesting people facing the same challenges and thriving for innovative solutions. My colleagues at KEK-CDC had similar questions. I hope that through this short overview and analysis of adaptions of evaluation processes, I can provide some kind of orientation and guidance. The guidelines will be presented to and validated with the team of KEK-CDC at an internal event at the end of 2020.

However, I am conscious of the limitations of the scope of this short thesis on one hand and the fast pace of changes and innovations on the other hand. Therefore, the elaborated guideline shall be considered a snapshot of the current situation which needs updates and additions based on future lessons learnt and best practices. The experience of the evaluation hackathon and group work during the KEK-CDC retreat showed the diversity of needs and questions, all longing to be answered. Although a combination of literature review and interviews was not encouraged within the small scope of this thesis, I believe it added another layer of meaning and grounded the insights and arguments of scientific contributions and blog articles within my work context. While I tried my best to live up to it, there was the difficulty of covering the different steps of an evaluation process, while not losing the way in details and the variety of opportunities and alternatives.

My thanks go to my colleagues at KEK-CDC – especially to Magali Bernard for her valuable inputs – to Dr. Stefanie Krapp, Prof. Dr. Lars Balzer and Mrs. Marielle Schaer-Selby for their support, motivation and guidance, my learning group for the possibility to exchange and especially to Team Phoenix for the exhausting but thrilling work during the EvalHack.

# Bibliography


Ahsan, Sairana. 2020. «The perks of remote documentation». https://www.authoraid.info/en/news/details/1385/ (26. August 2020).

Balzer, Lars, and Wolfgang Beywl. 2018. *Evaluiert : erweitertes Planungsbuch für Evaluationen im Bildungsbereich*. 2. überarbeitete Auflage. Bern: hep der Bildungsverlag.

Bamberger, Michael, and Linda Mabry. 2020. *Realworld evaluation : working under budget, time, data, and political constraints*. Third edition. Los Angeles: SAGE.

Barratt, Monica J, and Alexia Maddox. 2016. «Active Engagement with Stigmatised Communities through Digital Ethnography». *Qualitative Research* 16(6): 701–19.

Beywl, Wolfgang. 2007. *Evaluation Schritt für Schritt : Planung von Evaluationen*. 1. Aufl. Darmstadt: Hiba.

Beywl, Wolfgang, and Melanie Niestroj. 2009. *Das A-B-C der wirkungsorientierten Evaluation : Glossar - Deutsch/Englisch - der wirkungsorientierten Evaluation*. 2., vollst. bearb. und erg. Aufl. Köln: Univation - Inst. für Evaluation.

Bezerra, Fabio. 2020. *Guide on alternative approaches in data and evidence generation during the Covid-19 crisis*. Unicef Mozambique. https://sinapse.gife.org.br/download/guide-on-alternative-approaches-in-data-and-evidence-generation-during-the-covid-19-crisis (11. September 2020).

Champion, Dean J. 2006. *Research methods for criminal justice and criminology*. 3rd ed. Upper Saddle River, N.J.: Pearson.

Chen, Julienne, and Pearlyn Neo. 2019. «Texting the Waters: An Assessment of Focus Groups Conducted via the WhatsApp Smartphone Messaging Application». *Methodological Innovations* 12(3): 205979911988427.

Cronbach, Lee Joseph, and Karen Shapiro. 1982. *Designing evaluations of educational and social programs*. San Francisco [etc.]: Jossey-Bass.

DEval/AK-Epol-HuHi. 2020. «Evidenzbasierte internationale Zusammenarbeit in Zeiten der Corona-Pandemie. Auswirkungen von COVID-19 auf Evaluierungen in der EZ und humanitären Hilfe».

Dozois, Elizabeth u. a. 2011. *DE 201: A Practitioner's Guide to Developmental Evaluation*. Montréal, Que.: J.W. McConnell Family Foundation. https://www.deslibris.ca/ID/226187 (25. September 2020).

ESS, hrsg. 2019. «Evaluation in Hard-to-reach areas»: https://europa.eu/capacity4dev/devco-ess/documents/evaluation-hard-reach-areas-final-call-action-document-lessons-2019-cycle-conferences (2. July 2020).

European Evaluation. 2020. «Evaluation in Times of COVID-19». *European Evaluation*. https://europeanevaluation.org/evaluation-in-times-of-covid-19/ (30. August 2020).

FAO. 2020. «Risk Analysis and Guidance for the Management and Conduct of Evaluations during International and National Level COVID-19 Crisis and Restrictions».

Forrestal, Sarah G., Angela Valdovinos D'Angelo, and Lisa Klein Vogel. 2015. «Considerations for and Lessons Learned from Online, Synchronous Focus Groups». *Survey Practice* 8(3): 2844.

Garcia, Coline. 2020. *Covid-19 Crisis: HOW TO ADAPT YOUR DATA COLLECTION FOR MONITORING AND ACCOUNTABILITY?* CartONG.

Garcia, Jeneen R. 2020. «When Life Gives You Lemons…or a Pandemic - Finding Opportunity in Disaster». *Earth-Eval*. https://www.eartheval.org/blog/opportunity-disaster-2020-covid19 (5. September 2020).

Hagan, Frank E. 2018. *Research methods in criminal justice and criminology*. Tenth edition. New York, NY: Pearson.

Hinchcliffe, Vanessa, and Helen Gavin. 2009. «Social and Virtual Networks: Evaluating Synchronous Online Interviewing Using Instant Messenger»: 25.

IEO/UNDP. 2020. «Good practises under Covid-19: Joint Guidance Note for Evaluation Units». http://web.undp.org/evaluation/guideline/documents/covid19/IEOOECD_DAC_Joint-Guidance_COVID19.pdf (13. September 2020).

IFAD. 2020. «IOE Continues Implementing Good Evaluation and Data Collection Practices during the COVID-19 Outbreak». *IFAD*. https://www.ifad.org/en/web/ioe/news-details/asset/41896433 (26. September 2020).

ILO. 2020. «Implications of COVID-19 on evaluations in the ILO: Practical tips on adapting to the situation». https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_744068.pdf. (13. September 2020).

Janghorban, Roksana, Robab Latifnejad Roudsari, and Ali Taghipour. 2014. «Skype Interviewing: The New Generation of Online Synchronous Interview in Qualitative Research». *International Journal of Qualitative Studies on Health and Well-being* 9(1): 24152.

Josephson, Leah et al. 2020. *How Can We Use Evaluation to Support Decision-Making and Reflection in This Time of Community Crisis?* https://www.betterevaluation.org/en/resources/how-can-we-use-evaluation-support-decision-making-and-reflection-time-community-crisis (26. September 2020).

Jowett, Adam. 2020. «Carrying out Qualitative Research under Lockdown – Practical and Ethical Considerations». *Impact of Social Sciences*. https://blogs.lse.ac.uk/impactofsocialsciences/2020/04/20/carrying-out-qualitative-research-under-lockdown-practical-and-ethical-considerations/ (26. September 2020).

Krouwel, Matthew, Kate Jolly, and Sheila Greenfield. 2019. «Comparing Skype (Video Calling) and in-Person Qualitative Interview Modes in a Study of People with Irritable Bowel Syndrome – an Exploratory Comparative Analysis». *BMC Medical Research Methodology* 19(1): 219.

Lune, Howard, and Bruce L. Berg. 2017. *Qualitative research methods for the social sciences*. Ninth edition, global edition / Howard Lune, Bruce L. Berg. Harlow: Pearson.

Lysy, Chris. 2020. «The Evaluation Mindset: Evaluation in a Crisis». *Freshspectrum*. https://freshspectrum.com/the-evaluation-mindset-evaluation-in-a-crisis/ (2. September 2020).

Macfarlan, Alice. 2020a. «Adapting Evaluation in the Time of COVID-19 - Part 1: MANAGE». *BetterEvaluation*. https://www.betterevaluation.org/en/blog/adapting-evaluation-time-covid-19-part-1-manage (30. August 2020).

———. 2020b. «Adapting Evaluation in the Time of COVID-19 — Part 3: FRAME». *BetterEvaluation*. https://www.betterevaluation.org/en/blog/adapting-evaluation-time-covid-19-part-3-frame (30. August 2020).

Maloney, Jade. 2020. «Evaluating in a Pandemic: Why and How and When?» *AES*. https://www.aes.asn.au/aes-blog/evaluating-in-a-pandemic-why-and-how-and-when (25. September 2020).

OECD/DAC Network on Development Evaluation. 2019. «Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use».

https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf (13. September 2020).

Ofir, Zenda. 2020. «Actions for evaluators in and beyond the time of COVID-19». *Evaluation for Development*. https://zendaofir.com/wp-content/uploads/2020/04/Action-List-1200px.png (13. September 2020).

Opdenakker, Raymond. 2006. «Advantages and Disadvantages of Four Interview Techniques in Qualitative Research». *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 7(4). https://www.qualitative-research.net/index.php/fqs/article/view/175 (26. August 2020).

Open Access Repositories. 2020. «Coronavirus May Be Encouraging Publishers to Pursue Open Access». https://www.insidehighered.com/news/2020/05/15/coronavirus-may-be-encouraging-publishers-pursue-open-access (26. September 2020).

Parker, Cynthia Silva. 2020. «Living (and Working Virtually) in Uncertainty». *Interaction Institute for Social Change*. https://interactioninstitute.org/living-and-working-virtually-in-uncertainty/ (4. September 2020).

Patton, Michael Quinn. 2011. *Developmental evaluation : applying complexity concepts to enhance innovation and use*. New York: Guilford Press.

———. 2020. «Evaluation Implications of the Coronavirus Global Health Pandemic Emergency». *Blue Marble Evaluation*. https://bluemarbleeval.org/latest/evaluation-implications-coronavirus-global-health-pandemic-emergency (30. August 2020).

Petersson, Gustav Jakob, and Jonathan D. Breul, hrsg. 2018. *Cyber society, big data, and evaluation*. First issued in paperback. London: Routledge, Taylor & Francis Group.

Raab, Michaela. 2020. «International Evaluation in Times of the New Coronavirus». *people-centred development*. https://www.developblog.org/2020/03/international-evaluation-in-times-of.html (2. September 2020).

Raimondo, Estelle, Jos Vaessen, and Mariana Branco. 2020. «Adapting Evaluation Designs in Times of COVID-19 (Coronavirus): Four Questions to Guide Decisions». https://ieg.worldbankgroup.org/blog/adapting-evaluation-designs-times-covid-19-coronavirus-four-questions-guide-decisions (13. September 2020).

Ritz, Adrian. 2003. *Evaluation von New Public Management : Grundlagen und empirische Ergebnisse einer Reform-Evaluation aus der schweizerischen Bundesverwaltung*. [S.l.]: [s.n.].

SEVAL. 2017. «Erläuterungen zu den Evaluationsstandards der Schweizerischen Evaluationsgesellschaft». https://www.seval.ch/app/uploads/2018/01/SEVAL-Standards_2016-Erl%C3%A4uterungen.pdf (13. September 2020).

Sullivan, Jessica R. 2012. «Skype: An Appropriate Method of Data Collection for Qualitative Interviews?» : 8.

Ugarte, Rodrigo. 2020. «The Covid-19 Opportunity: Creating More Ethical and Sustainable Research Practices». *Items*. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/the-covid-19-opportunity-creating-more-ethical-and-sustainable-research-practices/ (13. September 2020).

UN Women. 2020. «Pocket tool for managing evaluation during the COVID-19 pandemic». https://europa.eu/capacity4dev/devco-ess/documents/pocket-tool-gender-responsive-evaluation-and-research-during-covid-19 (13. September 2020).

UNFPA Evaluation Office. 2020. «Adapting evaluations to the COVID-19 pandemic». https://www.unfpa.org/sites/default/files/admin-resource/FINAL_Covid19_and_Eval.pdf (13. September 2020).

Vaessen, Jos, and Estelle Raimondo. 2020. «Conducting Evaluations in Times of COVID-19 (Coronavirus)». https://ieg.worldbank.org/blog/conducting-evaluations-times-covid-19-coronavirus (2. Juli 2020).

Vandercasteelen, Joachim. 2020. «When Evaluators Cannot Make It to the Field, They Can Always Observe from Space». https://ieg.worldbank.org/blog/when-evaluators-cannot-make-it-field-they-can-always-observe-space (8. August 2020).

Weitzenegger, Karsten. 2020. «Evaluation in the COVID-19 Pandemic Mode, Part 1». *Monitoring & Evaluation & Learning*. https://deveval.wordpress.com/2020/05/07/evaluation-in-the-covid-19-pandemic-mode-part-1/ (30. August 2020).

York, Pete, and Michael Bamberger. 2020. *Measuring Results and Impact in the Age of Big Data: The Nexus of Evaluation, Analytics, and Digital Technology*. The Rockefeller Foundation. https://www.rockefellerfoundation.org/report/measuring-results-and-impact-in-the-age-of-big-data-the-nexus-of-evaluation-analytics-and-digital-technology/ (11. September 2020).

# Annex 1: Abbreviations

| | |
|---|---|
| AEA | American Evaluation Association |
| AES | Australian Evaluation Society |
| DAC | Development Assistance Committee |
| DeGEval | German Evaluation Society |
| ESS | Evaluation Support Service to the European Commission |
| F2f | Face-to-face |
| IFAD | International Fund for Agricultural Development |
| IEG | Independent Evaluation Group |
| ILO | International Labour Organization |
| IOE | Independent Office of Evaluation |
| IPDET | International program for development evaluation training |
| NGO | Non-governmental organization |
| OECD | Organization for Economic Co-operation and Development |
| SEVAL | Swiss Evaluation Society |
| ToR | Terms of Reference |
| UNFPA | United Nations Population Funds |
| VOIP | Voice over the Internet Protocol |

# Annex 2: Index of tables and figures

**Tables**



**Figures**

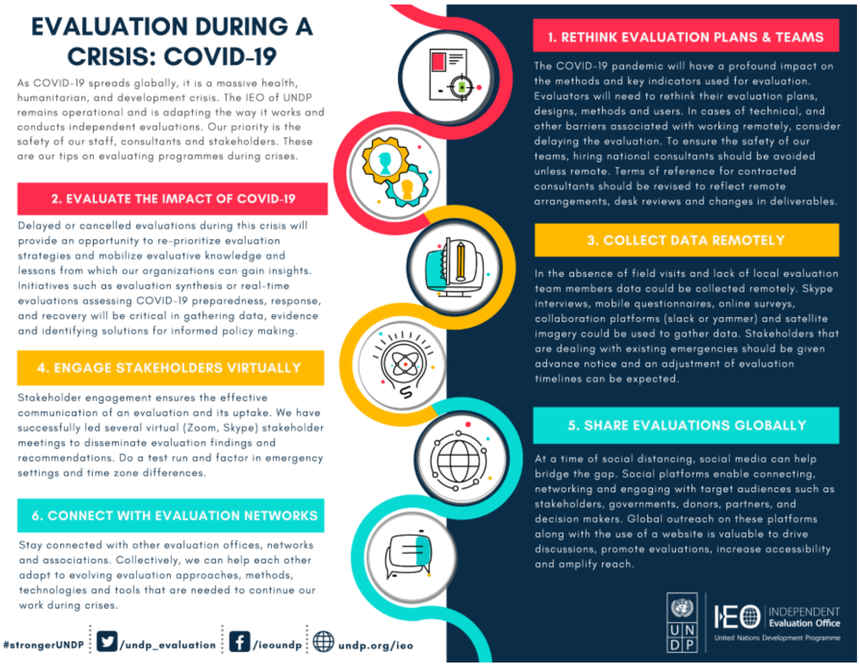
## Annex 3: Glossary of evaluation terms

The following terms and translations are based on Beywl/Niestroj (2009).

| German | English |
|---|---|
| Auftraggebende (der Evaluation) | clients; commissioners (of an evaluation) |
| Beteiligte und Betroffene | active and passive stakeholders |
| Datenerhebung | data collection |
| Ergebnisvermittlung | communication of findings |
| Evaluationsplan | evaluation design |
| Fragestellungen | evaluation questions |
| Gegenstand | evaluand; evaluation object |
| Gegenstandsbestimmung | identification of the evaluand |
| Informationsgewinnung | information acquisition |
| Zweck | purpose |

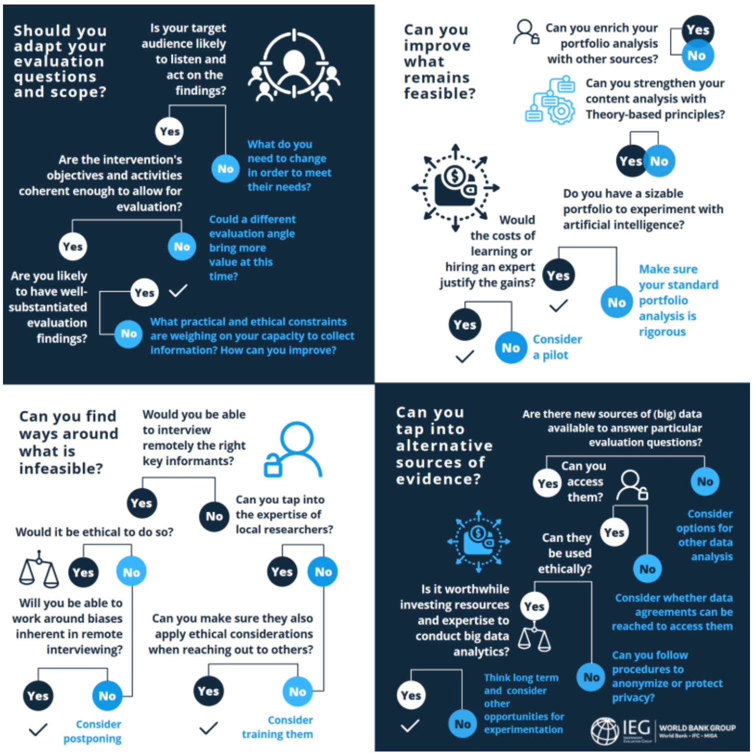
## Annex 4: Illustrations and infographics

*Figure 2: Evaluation during a crisis (UNDP 2020)*




*http://web.undp.org/evaluation/documents/infographics/Evaluation-during_crisis-COVID19.pdf*

*Figure 3: Covid Evaluation Decision Tree (IEG 2020)*




*http://ieg.worldbank.org/sites/default/files/Data/Blog-images/Covid_Eval_DecisionTree.pdf*

*Figure 4: Digital data collection during #Covid-19 (IEG 2020)*

*Table 7: Important aspects of online interviewing methods (own compilation)*

| | |
|---|---|
| E-Mail | • Asynchronous<br>• Allows for reflection by participant and interviewer, there is no need to agree on a common time slot and answers are directly documented. No strong internet connection required.<br>• Reduced spontaneity and commitment by interviewee, lack of body language and facial expressions |
| Instant Messenger | • Asynchronous or synchronous<br>• Faster and more conversational pace than e-mail and answers are also documented. No strong internet connectivity required.<br>• Reduced spontaneity as well as lack of body language and facial expressions. |
| Voice messages | • Asynchronous or synchronous<br>• Low internet connection is sufficient, answers are recorded and documented.<br>• Voice adds a layer of proximity. Nevertheless, the spontaneity of answers may also be reduced and the lack of body language and facial expressions remains. |
| Phone call | • Synchronous<br>• Common time slot needs to be found and answers are not documented.<br>• Possibility to notice spontaneous reactions and moods. Having no visual aspect can be seen both as a disadvantage or an asset, as it might eliminate layers of expectatuons and prejudices. |
| Video call | • Synchronous<br>• Experiece of seeing and hearing the interviewee, closest to a in-person f2f interview while being geographically seperate.<br>• Possibility to record interview, but transcription necessary. |

*Figure 5: Best practices for planning and implementing online, synchronous focus groups (Forrestal et al. 2015, 4)*

## Annex 5: Submission of challenge to the IPDET Hackathon

*Table 8: Submission of challenge to the IPDET evaluation hackathon*

| **Using technology for "off-site" Evaluations** | |
|---|---|
| **First Name/Country** | **Last Name/Organization** |
| Marina | Häusermann |
| Switzerland | KEK-CDC Consultants |
| Email | haeusermann@kek.ch |
| Category: | methodological |
| Due to the situation caused by the Covid-19 pandemic worldwide, travels abroad are currently and during the coming months impossible. Therefore, on site evaluations cannot be carried out and we have to look for different means to gather the necessary data, e.g. with the help of technological means. Similar conditions for an evaluation could also arise for other reasons. Be it due to a critical security situation in the destination country or budgetary or ecological considerations, on the basis of which must be negotiated and assessed which parts of the evaluation could also be carried out "off-site" instead of on site. | |
| Given the current situation, it is important to consider which projects/programmes/activities can be evaluated from a distance and which evaluation approaches and methods are particularly suitable for this purpose. Due to the economic challenges triggered by the Corona crisis, it will soon be all the more important to be able to prove that expenditure, e.g. for development cooperation, is spent effectively and efficiently. This will require creativity in adapting conventional monitoring and evaluation systems, as Jeff Chelsky and Lauren Kelly wrote in a recent blog article of the World Bank's Independent Evaluation Group. While evaluators worldwide have been discussing about instruments that can be used in evaluations without field access in blog articles for years, they only tackle single issues (e.g. interviews). In regard of the challenges ahead of us, there should be more consideration given to how we can use modern technology to gather data or make use of data, which has already been gathered by other means. | |
| How can we use modern technologies and existing data in order to carry out "off-site" evaluations without access to the field, replacing methods of data gathering by others? What potential, but also what lacks does such an approach have? | |

## Annex 6: Documentation of group work at the KEK-CDC Retreat

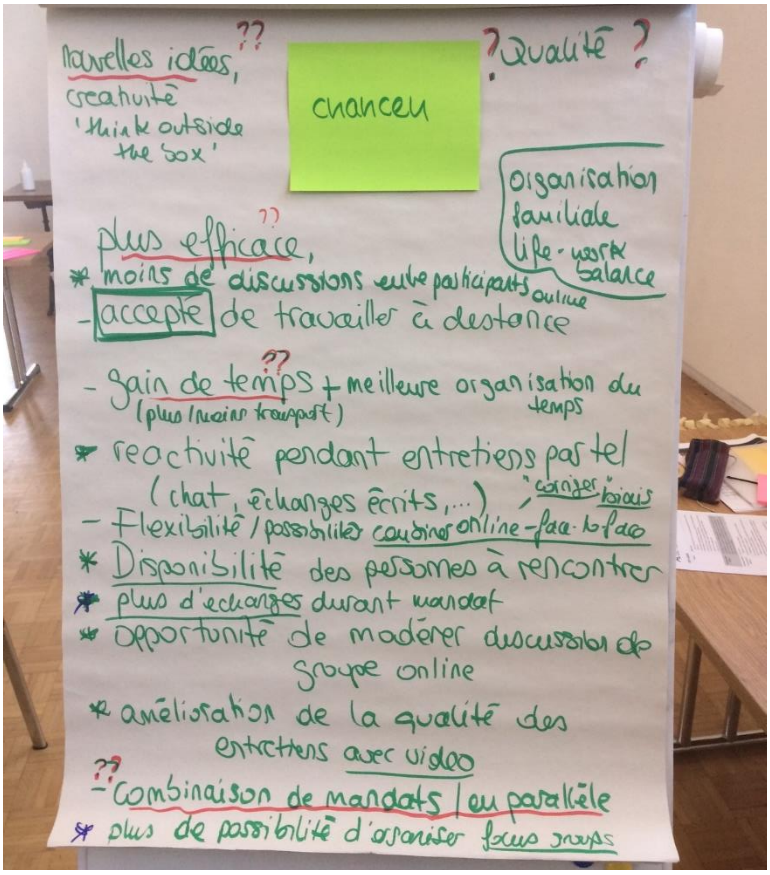



*Table 9: Challenges and constraints identified by KEK-CDC*

| | |
|---|---|
| *Planning* | • Different time zones<br>• Thought of "gaining" time while not travelling. Nevertheless debriefings and exchanges have to be scheduled, as there are no "natural opportunities" for them as in the field (in the car, at dinner, etc.) |
| *Infrastructure & Skills* | • Lack of technological infrastructure<br>• Lack of necessary skills to use digital tools properly<br>• Difficulty to find trustworthy, professional, and available local consultant |
| *Stakeholder engagement* | • Stakeholders might not be available / interested anymore<br>• Difficult to set up trust / no feeling of proximity / Building rapport with stakeholders impossible<br>• Lack of group dynamic, body language etc.<br>• Network formation and creative group processes more complicated<br>• Difficulty of hybrid formats and settings with big groups<br>• Selection bias due to remote data collection |
| *Data collection* | • No "feeling" of the context; lack of observations<br>• Difficulty to confirm validity of data<br>• Lack of "natural opportunities" for informal chats, personal contact, exchanges and de-briefings<br>• Complementation with secondary/big data as asset. But lack of skills and previous experience.<br>• Less discussions online than f2f |
| *Risks* | • Risks of overloads -> plan breaks<br>• Limited flexibility (e.g. interview partners predefined) and independence<br>• Limitations for formative evaluations, as development of the situation remains uncertain |

*Table 10: Opportunities identified by KEK-CDC*

| | |
|---|---|
| *Set-up* | • Working remotely / from home is widely accepted<br>• Balance between private/family life and work might be easier<br>• Possibilities to think outside the box<br>• Easier to work in different mandates at the same time |
| *Planning* | • Better planning due to less mobility<br>• More flexibility to schedule interviews and focus group discussions<br>• More and several exchanges with stakeholders possible, as the time frame is longer |

## Selbstständigkeitserklärung

Ich erkläre hiermit, dass ich die Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen (inkl. elektronischen Quellen) benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen sind, habe ich als solche gekennzeichnet. Die Mitwirkung Dritter an der Konzeption, Durchführung und Schriftfassung der Arbeit habe ich in vollem Umfang offen gelegt. Mir ist bekannt, dass andernfalls die Arbeit mit der Note 1 bewertet wird und der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes über die Universität vom 5. September 1996 und Artikel 20 des Universitätsstatuts vom 17. Dezember 1997 zum Entzug des aufgrund dieser Arbeit verliehenen Abschlusses berechtigt ist.