ERiK report also available via **www.dji.de/erik**

ERiK

Methodological Report V

Design and Implementation of the ERiK-Surveys 2024

Forschungsverbund
tu DJI
Deutsches Jugendinstitut
Technische Universität Dortmund

DJI
Deutsches Jugendinstitut

wbv

ERiK-Methodological Report V

Design and Implementation of the ERiK-Surveys 2024

ERiK

Methodological Report V

Design and Implementation of the ERiK-Surveys 2024

Forschungsverbund
tu DJI
Deutsches Jugendinstitut
Technische Universität Dortmund

DJI
Deutsches
Jugendinstitut

wbv

Researching children, youth and families at the intersection of science, policy, and professional practice

The German Youth Institute (DJI) is one of the largest social science research institutes in Europe with an experience of over 60 years. The DJI conducts empirical studies into the life situations of children, young people and families, and provides policy advice to the German national government, the German federal states and local authorities as well as key impulses for professional practice.

Founded in 1963, the governing body of the institute is a non-profit association with members from the fields of politics and science as well as from child, youth and family welfare institutions and organisations. The DJI receives the majority of its funding from the Federal Ministry for Education, Family Affairs, Senior Citizens, Women and Youth, and the German federal states. The Federal Ministry of Education and Research, the European Commission as well as foundations and other science funding institutions provide additional funding.

Currently, about 380 staff members work and conduct research at the two locations in Munich and Halle (Saale).

Authors:
Franz Classe, Jakob J. Gilg, Caterina L. Sanchez Steinhagen, Sonja Herrmann, Susanne Kuger

Design and layout of inner section:
Christiane Zay, Passau

Interior typesetting:
paginamedia

Editing:
Baker & Company, Munich

2026 wbv publication
a division of
wbv Media GmbH & Co. KG, Bielefeld

Overall production:
wbv Media, Bielefeld,
Auf dem Esch 4, 33619 Bielefeld,
service@wbv.de
wbv.de

Cover design:
Christiane Zay, Passau

Order number: I79967
ISBN Print: 9783763979967
ISBN E-Book: 9783763979974
DOI: 10.3278/9783763979974

Printed in Germany

This publication is freely available for download at
wbv-open-access.de

This publication, with the exception of the cover photo, is published under the following Creative Commons:
creativecommons.org/licenses/by-sa/4.0/en

All trademarks, trade names, logos, or company names referenced in this publication are used for identification only and are the property of their respective owners

In accordance with Section 44b of the German Copyright Act (UrhG) the publisher reserves the right to engage in text and data mining, which is hereby prohibited to third parties without the publisher's consent.

Bibliographic information of the German National Library
The German National Library lists this publication in the German National Bibliography; detailed bibliographic data is available on the Internet at http://dnb.d-nb.de.

Funded by:

Federal Ministry
for Education, Family Affairs, Senior Citizens,
Women and Youth

Contents

1. Preface 7

2. Introduction 9

3. Target Populations 11
   - 3.1 Directors of Day-Care Centres 11
   - 3.2 Pedagogical Staff of Day-Care Centres 11
   - 3.3 Family Day-Care Workers 11
   - 3.4 Youth Welfare Offices 12
   - 3.5 Providers of Day-Care Centres 12
   - 3.6 Parents of Children in Day-Care 12

4. Sampling Frames 13
   - 4.1 Federal State Lists 13
   - 4.2 Youth Welfare Offices & Family Day-Care Workers 15

5. Sampling Designs 17
   - 5.1 Day-Care Centres: Directors & Pedagogical Staff 17
   - 5.2 Youth Welfare Offices & Family Day-Care Workers 18
   - 5.3 Providers 19

6. ERiK Parent Survey (KiBS) 21

7. Implementation of the ERiK-Surveys 2024 23
   - 7.1 Questionnaires 23
   - 7.2 Field Period 23
   - 7.3 Providers 28
   - 7.4 Number of Completed Questionnaires and Response Rates 29
   - 7.5 Nonresponse Surveys 31

8. Data Quality of the ERiK-Surveys 2024 33
   - 8.1 Limited Comparability of the 2020, 2022 and 2024 Surveys 33
   - 8.2 Weighting 34
   - 8.3 Informative Value at Federal and State Level 38

9. Summary 41

10. Bibliography 43

11. Appendix 45

1. Preface

The ERiK Methodological Report IV (German: ERiK-Methodenbericht IV) is the fifth methodological report in the study "An indicator-based monitoring of structural quality in the German early childhood education and care system" (German: Entwicklung von Rahmenbedingungen in der Kindertagesbetreuung – indikatorengestützte Qualitätsbeobachtung – ERiK). ERiK is a nationwide study that aims to provide data-based knowledge on the quality of early childhood education and care (ECEC) in Germany. It includes regular surveys of parents, directors of day-care centres, pedagogical staff, family day-care workers, youth welfare offices and providers of day-care centres. The DJI conducted the first surveys in 2020, followed by further surveys in 2022 and 2024.

While the methodological aspects of the ERiK-Surveys 2022 are presented in two separate reports (Gilg et al., 2023; Herrmann et al., 2024), the present report covers all of the methodological background information regarding the ERiK-Surveys 2024. Also, this report was prepared with the assistance of artificial intelligence (AI) tools for editorial refinement. All content was critically reviewed, verified, and finalized by the authors. The information presented in this report is the result of the collective efforts of the entire ERiK team that actively contributed to the ERiK-Surveys 2024 at the German Youth Institute (Deutsches Jugendinstitut – DJI). Special acknowledgement is extended to the members of the ERiK steering group, namely Prof. Dr. Susanne Kuger, Prof. Dr. Bernhard Kalicki, Dr. Christiane Meiner-Teubner, Dr. Sina Fackler, Dr. Melanie Böwing-Schmalenbrock and Christine Bopp. Additionally, gratitude is expressed to the ERiK team colleagues for their invaluable feedback, including Janette Buchmann, Yasmin Öztürk, Melina Preuß, Johanna Romefort, Anna Selmayr, Nadira Tursun, and Lisa Ulrich.

The monitoring project contributes to an ongoing effort to advance the quality of early childhood education and care in Germany. As a result, this report has been enriched by the various discussions held with colleagues of the Federal Ministry for Education, Family Affairs, Senior Citizens, Women, and Youth (BMBFSFJ), representatives from the ministries of the federal states, and experts from other fields. Engaging with different committees and stakeholders has been instrumental in gathering central technical insights for the report. In addition, we would like to thank Dr. Ludovica Gambaro of the Federal Institute for Population Research (BiB) for reviewing the report and her helpful suggestions for improvement. Moreover, we would like to express our gratitude to the infas Institute for Applied Social Research and Communication for their dedication and efforts in conducting the ERiK-Surveys 2024.

Lastly, the report has been enriched by the invaluable practical assistance of numerous individuals. We extend our special appreciation to Isabel Becker, Eugenia Zimmermann, Cora Szabo, and the dedicated student assistants within the ERiK team.

Munich, February 26

Dr. Franz L. Classe – *Co-Head of the ERiK Project, DJI*
Jakob J. Gilg – *Research Associate within the Department of Social Monitoring and Methodology, DJI*
Caterina L. Sanchez Steinhagen – *Research Associate within the Department of Social Monitoring and Methodology, DJI*

2. Introduction

The Act on the Further Development of Quality and the Improvement of Participation in Day Care Facilities and in Child Day Care (German abbrev.: KiQuTG) came into force on 1 January 2019 with the aim of improving the early childhood education and care (ECEC) system throughout Germany. The extension of the law with a stronger focus on refining the quality of childcare passed into law on 1 January 2023. Since then, the law has been abbreviated in German as the "KitTa-Qualitätsgesetz" or "Childcare Quality Act". The law was extended again in 2024 and came into force on 1 January 2025. The (further) development of the law is linked to indicator-based monitoring, which is being implemented as part of the ERiK ("indicator-based monitoring of structural quality in the German early childhood education and care system") project. The aim is to use research reports published every two years to empirically observe and classify developments in quality and participation in the early childhood education and care (ECEC) system in Germany. In this context, reliable data on the structure and conditions of ECEC are essential for monitoring developments in this area over time.

The ERiK project has been conducted since 2020 with the goal of generating systematic insights into the nationwide quality of ECEC. In 2024, the ERiK-Surveys were carried out for the third time, continuing the established monitoring framework. Further surveys are planned for 2026, targeting the same populations as in 2024.

Its target populations include directors, pedagogical staff, and providers of day-care centres, youth welfare offices, and family day-care workers in all German federal states. In addition, parents of children not yet attending school were surveyed within the context of the DJI Child Care Study (KiBS; Kinderbetreuungsstudie). The approach of surveying different populations aims at ensuring that diverse perspectives of actors within the German ECEC system are adequately reflected in the data and consequently in subsequent analyses and reporting. This aspect of multi-perspectivity is a central feature of the ERiK project and contributes to a comprehensive understanding of the ECEC landscape in Germany.

This methodological report provides an overview of the design and implementation of the ERiK-Surveys 2024. Specifically, it contains information on:

- Target populations
- Sampling frames and sampling designs
- The integration of KiBS
- Survey implementation (including fieldwork, response rates, and nonresponse analyses)
- Data quality and weighting procedures.

In contrast to the previous ERiK-Surveys (2020, 2022), which were documented in two separate volumes, the present report covers all methodological aspects of the ERiK-Surveys 2024 in one single publication. Overall, this report provides a comprehensive overview of the methodology of the ERiK-Surveys 2024. Furthermore, it offers insights into data quality at both state and federal levels, as well as the comparability of results across survey years.

3. Target Populations

In line with the principle of multiperspectivity, the ERiK-Surveys cover different actors in the field of ECEC. As in the 2020 and 2022 surveys, the target populations in 2024 are directors and pedagogical staff of day-care centres, family day-care workers, youth welfare offices, providers of day-care centres, and parents. The following sections briefly describe the target populations (for greater detail see ERiK Methodological Report III (Gilg et al., 2023)). Unless otherwise noted, population sizes (N) are based on the most recent KJH statistics at the time of surveys (Statistisches Bundesamt [Destatis], 2023).

3.1 Directors of Day-Care Centres

Directors of day-care centres are responsible for managerial tasks such as overseeing compliance with regulations, putting provider policies into practice, and supervising pedagogical staff. Moreover, depending on their employment agreements, they may also directly engage in pedagogical work with children (Buchmann et al., 2022).

In ERiK, the population of day-care centre directors is defined as employees of day-care centres that are responsible for the majority of managerial tasks in the centre. In the case of management teams, where managerial tasks are split equally between two or more people, the employee with the most recent birthday was selected to participate (see Gilg et al., 2023), ensuring there is exactly one director per day-care centre. The size of this target population is estimated at N = 56,077 (Destatis, 2023).

3.2 Pedagogical Staff of Day-Care Centres

In ERiK, pedagogical staff are defined as full- or part-time employees who work pedagogically with children in day-care centres for children not yet attending school. This definition includes employees providing educational and care services while still in training and excludes volunteers as well as non-pedagogical staff, such as janitors and kitchen staff.

In 2020, pedagogical staff with managerial responsibilities were included in the target population. However, in 2022, they were explicitly excluded in order to separate the perspectives of directors and staff more clearly (see Gilg et al., 2023). This change meant that calibration with the KJH statistics was no longer possible (see Fackler et al., 2026). Staff with managerial responsibilities were re-included in 2024 (see Section 8.1 for more information). The described target population consists of N = 648,514 employees of day-care centres (Destatis, 2023).

3.3 Family Day-Care Workers

The target population of family day-care workers is defined as people that hold a valid permit issued by the local youth welfare office, are publicly funded, and currently provide ECEC according to § 22 and § 23 SGB VIII for at least one child not yet attending school (for more details see Gilg et al., 2023). Compared to the KJH statistics, which define the population based on public funding according to § 23 SGB VIII, the ERiK definition relies on possession of a valid permit. It remains unclear to what extent these approaches capture different populations, as information on overlap is missing (see Gilg

et al., 2023). The size of the target population for family day-care workers with day-care workers caring exclusively for schoolchildren excluded is estimated at N = 41,233 (based on Destatis, 2023).

3.4 Youth Welfare Offices

The ERiK target population of youth welfare offices comprises all local youth offices in Germany, except for the seven youth offices of city districts in Hamburg that do not have the same role as the other local youth offices in Germany. Instead, the state youth office fulfils the tasks relevant for inclusion in the ERiK target population, that is the assignment of services (SGB VIII) and granting operating permits for family day-care workers (Gilg et al., 2023). By contrast, the twelve district youth offices in Berlin are each included separately, as they operate independently from each other (see Gilg et al., 2023). Taking these cases into account results in a target population of 567 youth welfare offices that should complete the ERiK questionnaire.

3.5 Providers of Day-Care Centres

The definition of the population of providers of childcare has not changed since the first ERiK Survey in 2020. All types of providers that hold a permit according to § 45 SGB VIII and whose day-care centres (also) care for children not yet attending school are part of the target population (Gilg et al., 2023). As is the case for the previous ERiK-Surveys, the size of this population is unknown, since providers are not reliably recorded in official statistics. As the ERiK sampling frame contains 20,665 providers (see Section 4.1), we assume this to be the ERiK target population for the survey.

3.6 Parents of Children in Day-Care

The survey of parents is conducted by including specific questions in the KiBS (German: Kinderbetreuungsstudie), a panel study that has surveyed parents since 2012. The ERiK target population is defined as the sub-group of the KiBS target population comprising all parents in Germany with at least one child not yet attending school. For additional information on KiBS, see Section 6 below and Wieschke et al. (2025).

4. Sampling Frames

One fundamental part of the sampling procedure is the acquisition of a suitable sampling frame that minimizes coverage error (Groves, 1989) by closely matching the target population. In the field of German ECEC, this is particularly challenging, as official registers for some relevant groups are not publicly accessible.

This chapter presents the sampling frames used for the ERiK-Surveys 2024. The following sections detail the frames for childcare providers and day-care centres, youth welfare offices, and family day-care workers.

4.1 Federal State Lists

Since commercially acquired lists for day-care centres and providers proved to be incomplete and inaccurate in the ERiK-Surveys 2020 (D. Schacht, 2022), official lists provided by the German federal states have been used to construct the sampling frames for day-care centres and providers since the ERiK-Surveys 2022. The current address lists were provided by the German federal states between March and September 2023. Together, they form a reliable basis, significantly improving upon the coverage errors associated with commercial lists used in the ERiK-Surveys 2020.

The federal states provided information on 60,848 day-care centres belonging to 21,249 providers. The lists provided contain information such as names and postal addresses for all entities. As Table 4.1-1 shows, the availability of additional information varied considerably. Specifically, information on the type of provider that allows distinguishing, for example, public providers from independent, non-charitable providers, was considered necessary for weighting. Likewise, information on type of centre was necessary to exclude centres catering exclusively for school children typically in the afternoon.

A key challenge in cleaning and constructing the sampling frame according to our target population definition was the identification and exclusion of day-care centres offering care exclusively for schoolchildren (in German: “reine Horte”). The approach to this problem was twofold and depended on the information provided by each state:

1. For states that supplied a „type of centre“ variable: Centres explicitly classified as a “Hort” or similar were directly excluded from the sampling frame.
2. For states that did not supply this variable: A manual filtering process based on the centre‘s name was necessary. Centres were excluded if their name contained terms like “Schulhort”, “Schülerhort”, or similar compounds of “Schule” and “Hort”. To avoid over-exclusion, a centre was retained if its name also contained an indicator of the target population (e.g., “Kindergarten”, “Kindertages”, or “Krippe”).

Table 4.1-1: **Additional information in the lists of names and addresses of day-care centres and provider, by federal state**

	Day-Care Centres			Providers		
	Email Address	Telephone Number	Type of Centre	Email Address	Telephone Number	Type of Provider
Baden-Wuerttemberg	–	X	X	X	X	X
Bavaria	–	X	–	–	X	X
Berlin	X	X	–	X	X	X
Brandenburg	X	X	X	X	X	X
Bremen	X	X	X	X	X	X
Hamburg	X	X	X	X	X	–
Hesse	X	X	X	X	X	X
Mecklenburg-West.Pom.	X	X	–	–	–	X
Lower Saxony	–	–	–	–	–	–
North Rhine-Westphalia	X	X	–	X	X	X
Rhineland-Palatinate	X	X	X	X	X	X
Saarland	X	X	–	X	X	X
Saxony	X	X	X	X	X	–
Saxony-Anhalt	X	X	X	X	X	X
Schleswig-Holstein	X	–	X	X	–	X
Thuringia	–	–	–	–	–	X

This cleaning process also involved the removal of duplicates (see Table 4.1-2).

Table 4.1-2: **Sampling Frames ERiK-Survey 2024 and 2022: Day-Care Centres**

	Total 2024	Proportion 2024	Total 2022	Proportion 2022
Number of day-care centres	60,848	100 %	60,078	100 %
Duplicate entries (and further exclusions)	117	0.19 %	72	0.12 %
Day-care centres exclusively for school children	3,702	6.08 %	2,483	4.13 %
Sampling Frame in 2024	57,029	93.72 %	57,523	95.75 %

After this, a data completion process was conducted in multiple stages to ensure the quality and completeness of the sampling frame. In a second stage, existing contact data and provider type information were taken from matching cases in the sampling frame of the ERiK-Surveys 2022 to fill corresponding gaps in the 2024 list. A key part of this stage was the standardization of the provider type variable across all federal states to ensure consistent classification. This necessary process, however, resulted in data loss, reducing the completeness of this variable from 82 % in the original lists to 61 %. For the third and final stage, the external research institute DataM was commissioned to research and impute all remaining empty entries for email addresses, telephone numbers, and provider types. During this stage, the existing contact information was also systematically checked for accuracy; entries that were identified as invalid or incorrect were subsequently deleted. This final imputation round successfully completed the provider type variable, achieving a completeness rate of 99 % and resulting in the final, verified sampling frame. The percentage of data completeness for the contact information variables at each stage of the process is documented for each federal state in Table 11-1 in the appendix.

The final sampling frames for the ERiK-Surveys 2024 comprised 57,029 day-care centres belonging to 20,665 providers. One significant advantage of these state-provided lists is their integrated structure, which allows for the precise merging of data collected from providers, centre directors, and pedagogical staff, forming a robust foundation for multilevel analysis.

4.2 Youth Welfare Offices & Family Day-Care Workers

The sampling strategy for the youth welfare offices (YWO) and family day-care workers (FDW) in the ERiK-Surveys 2024 remains interdependent, as no independent, official sampling frame exists for the latter target population.

The sampling frame for youth welfare offices is based on the established list used in the ERiK-Surveys 2020 and 2022. The list used in the ERiK-Surveys 2022, however, was based on the KJH statistics from 2020. This list was updated and verified for the ERiK-Surveys 2024 using the updated KJH statistics from 2023. The frame includes all relevant municipal youth welfare offices while excluding state-level offices (except the Hamburg state youth welfare office).

Two minor changes in the sampling frame have occurred since the last survey. The urban district of Eisenach, together with its municipal youth welfare office, was incorporated into the district of Wartburgkreis on 30 June 2021. In addition, the previously independent youth welfare office of the Villingen-Schwenningen municipal administration merged with the youth welfare office of the Schwarzwald-Baar-Kreis district administration on 1 July 2023. This reduced the total number of districts to 573. Furthermore, due to the decentralized organisation in Hamburg, the state youth welfare office participated in the survey on behalf of its seven district offices. Consequently, the number of eligible youth welfare offices for direct surveying (sampling frame YWO) was 567.

As there is no publicly available nationwide register of family day-care workers, the ERiK project continued to employ an indirect approach to reach this target population. Each family day-care worker in Germany is exclusively associated with a distinct youth welfare office. Therefore, the youth offices not only are surveyed in their own right, but also serve as conduit for the sampling frame of family day-care workers. According to the 2023 KJH statistics, there were 41,233 publicly funded family day-care workers across Germany associated with 573 youth welfare offices. However, according to the KJH statistics, in seven youth office districts, there were no registered family day-care workers. Thus, a total of 566 youth welfare offices were asked to forward the questionnaires to all family day-care workers in their respective districts. This includes the seven district youth welfare offices of Hamburg and excludes the state youth office of Hamburg.

Reaching family day-care providers through youth welfare offices is a necessary and reliable method, but nonetheless not entirely error-proof. For example, youth offices may inadvertently exclude some eligible individuals or include inactive ones (e.g. those on leave or retired). To mitigate this, the 2024 questionnaire includes screening questions to identify and later exclude ineligible respondents during data cleaning, ensuring the final sample aligns closely with the defined target population.

5. Sampling Designs

In the ERiK-Surveys 2024, we used different sampling designs for the different target populations. The designs were very similar to the ones employed in the ERiK-Surveys 2022 and only slightly adjusted. The designs' central aim was to sample enough cases to enable generalisable statements at the federal state level. Table 5-1 shows an overview of these designs. In the following sections, we present the sampling designs for the different populations in detail.

Table 5-1: **Sampling and Survey Design ERiK-Surveys 2024**

Target Population	Directors/ Centres	Pedagogical Staff	Family Day-Care Workers	Youth Welfare Offices	Providers of Centres
Population Size (KJH 2023)	57,605	641,817	41,233	567 / 573[1]	unknown
Sampling Frame	Address list of federal states	Indirect sampling	Indirect sampling	Self-generated address list	Address list of federal states
Sampling Design	Stratified sample of centres	Two-stage stratified cluste sample	Complete population survey	Complete population survey	Complete population survey
Gross Sample	17,678	41,396	41,233	567 / 573	20,665
Gross Sample: Paper and Online	14,295	33,560	18,885	567 / 573	0
Gross Sample: only Online	3,383	7,828	22,350	0	20,665
Target Net Sample Size	5,100	6,100	4,270	472	4,000

5.1 Day-Care Centres: Directors & Pedagogical Staff

This section first outlines the sampling design of the ERiK-Survey 2024 for directors and subsequently relates it to, and explains the corresponding design of, the pedagogical staff survey. As in the previous ERiK-Surveys, the sample of day-care centres is stratified across the German federal states and randomly sampled in each state independently.

The framework of the sampling design for directors remained largely unchanged compared to the ERiK-Surveys 2020 and 2022, with only minor adjustments (see Gilg et al., 2023). Specifically, the ERiK team updated the calculation of the gross sample by aligning the number of selected cases within each federal state with the most recently available child and youth welfare statistics (2023 KJH statistics). In addition, response rates from the ERiK-Surveys 2020 and 2022 were taken into account to adjust the sample sizes accordingly. These considerations led to a total gross sample of 17,678 centres that should be contacted to reach net samples large enough to make reliable state-

1 The number depends on how the youth offices in Hamburg are counted; i.e. if one wants the population size for the youth office survey (567) or if one counts the youth offices that should contact family day-care workers (573) (see Section 4.2).

ments about all federal states.[2] The gross sample size varied across federal states, ranging from 438 in Saarland to 1,622 in Berlin.

As in the ERiK-Surveys 2022, we gave 80 % of centres the option to choose between paper and online questionnaires, while 20 % only got a letter with the link to the online questionnaire (see also Table 5-1 above). The last row in the table shows the net sample sizes that we hoped to achieve during the field period.

Pedagogical staff were sampled indirectly in a two-stage process via the directors of day-care centres. This is analogous to the ERiK-Surveys 2022, where we asked the directors of day-care centres to distribute the survey materials to their pedagogical staff. As there is still no official list of pedagogical staff, there is no direct way to contact this target population. Additionally, this approach is aligned with the goal of multi-perspectivity because it allows us to obtain responses of directors and pedagogical staff from the same day-care centres.

For the number of questionnaires that directors should distribute to their pedagogical staff, we mostly followed the 2022 design (see Gilg et al., 2023). However, we changed the number of questionnaires that directors should forward to their pedagogical staff. Instead of one to six questionnaires per centre (depending on the size of the federal state), directors should distribute two to five questionnaires to their pedagogical staff. This reduces differences in clustering of pedagogical staff between federal states and minimises complexity in packaging the envelopes at the survey institute. The resulting gross sample consists of 41,396 pedagogical staff members that should be contacted by their respective directors. In instances where the number of pedagogical staff in a given centre exceeds the number that should be contacted according to our design, as is typically the case, the directors are requested to randomise the distribution by forwarding the questionnaires to the staff members whose birthdays occurred most recently.

In each centre, the survey mode options for directors and their pedagogical staff were kept identical in order to facilitate easier distribution of questionnaires by the directors. Consequently, 80 % of pedagogical staff were offered both paper and online questionnaire options with 20 % receiving online questionnaires only (see Table 5-1). In summary, the sampling design of the ERiK-Survey 2024 for pedagogical staff followed a stratified two-stage cluster approach, with day-care centres serving as the clusters and the federal states as the strata.

5.2 Youth Welfare Offices & Family Day-Care Workers

There were no changes in the sampling design of youth welfare offices compared to the previous ERiK-Surveys. As in 2020 and 2022, we conducted a complete population survey and asked all youth offices to complete a questionnaire. The limited total number of youth offices in Germany (and especially in some federal states) is problematic and necessitates a high response rate to enable us to make generalizable statements about the population. To facilitate this, we provided a paper and an online questionnaire to each of the 567 youth offices, allowing them to choose their preferred format.

The sampling design for family day-care workers is analogous to that of the ERiK-Survey 2022 and the design for pedagogical staff (see Gilg et al., 2023). Due to the absence of a complete sampling frame, the ERiK project employed an indirect sampling strategy by asking the youth offices3 to distribute questionnaires to all family day-care workers active in their district. This has two reasons: firstly, it reduces the time and effort required for distribution when no sampling or selection within the youth office is required. Secondly, our experience with this population shows that it is difficult to reach an adequate number of responses and that response rates (especially in some federal states)

2 Due to an error by the survey institute, five centres exclusively for school children were sampled in Bremen. After it was noticed, these cases were excluded and five substitute cases were drawn from the sampling frame in Bremen.

3 This includes the seven district youth offices in Hamburg, but not the state youth office in Hamburg.

tend to be low. By conducting a complete survey of the population, we hoped to remedy this somewhat and achieve acceptable net sample sizes. Additionally, we increased the proportion of family day-care workers that received both paper and online questionnaires from 3 % in 2022 to about 46 % by asking half of the youth offices to send out paper questionnaires. Having the option to choose the preferred questionnaire mode was expected to increase response rates (see e.g., Dykema et al., 2013; Olson et al., 2012).

As depicted in Table 5-1, these designs are both complete population surveys with gross samples of 567 youth offices (573 for the family day-care worker distribution) and 41,235 family day-care workers.

5.3 Providers

The sampling design for the ERiK-Survey of providers of childcare centres did not change substantially compared to the previous surveys (see D. Schacht, 2022; Gilg et al., 2023). As described in Section 4.1, the sampling frame consists of 20,665 cases, all of which we contacted, making this a complete population survey in all federal states. In the ERiK-Survey 2022, we had tested questionnaire modes and observed that the response rate among providers was not negatively impacted when they only received online questionnaires (Gilg et al., 2023). Thus, we sent each provider a letter with a link to the online questionnaire in the first contact. Only with the first reminder, providers that had not yet participated received a paper questionnaire as well.

6. ERiK Parent Survey (KiBS)

Since 2019, the ERiK parent survey has been conducted as part of the Childcare Study (KiBS; German: Kinderbetreuungsstudie), a yearly panel survey of parents in Germany conducted by the DJI on behalf of the Federal Ministry for Education, Family Affairs, Senior Citizens, Women and Youth (BMBFSFJ). In its current form, KiBS has collected data on day-care participation, childcare arrangements and needs, parental fees, satisfaction with childcare, as well as socio-structural characteristics (e.g., household composition and parental employment) since 2016. The KiBS target population consists of parents of children younger than 11 not yet attending a secondary school. The ERiK target population of parents is a sub-group of the KiBS target population, as only parents of children not yet attending primary are relevant for the monitoring of the Childcare Quality Act.

The KiBS sample consists of parents that had already participated in previous waves (panel) and refreshment participants. The proportion of panel participants is restricted to 50 % of the total sample to counteract panel selection effects (for more information see Wieschke et al., 2025). KiBS applies a two-stage stratified cluster design. First, 434 municipalities are randomly selected from the municipal registers of all German municipalities. Within these municipalities, children are randomly drawn from the residents' registers, and the corresponding parents invited to participate. Duplicates resulting from parents having more than one child are removed beforehand (for further detail see Wieschke et al., 2025).

In the ERiK-Surveys, filtering is applied directly in the questionnaire with the result that only parents of children not yet in school receive ERiK-specific questions. During data preparation, the sample is further restricted based on the child's reported school status and age, with implausible cases (regarding age) excluded. In the thirteenth wave of the survey (February – August 2024), this procedure resulted in a final sample size of N = 22,310 parents of children not yet attending school. Of these parents, 8,932 (40 %) are part of the panel sample, while 13,378 (60 %) come from the refreshment sample.

Because parents included in the ERiK-sample are selected through filtering KiBS participants and not independently it is not possible to calculate an ERiK-specific gross sample. To nonetheless derive the number of parents who could potentially participate in the ERiK-survey (i.e., an approximate gross sample) we calculated the numbers of children under six and under seven in the KiBS gross sample, as these two groups could potentially attend a day-care centre. We then used these two as alternative denominators to calculate a lower and upper bound of approximate response rates for the total, refreshment and panel samples. Table 6-1 presents the resulting upper and lower bounds for the total, refreshment, and panel sample response rates.

Table 6-1: **Upper and Lower Bounds of Response Rates**

Sample	Lower Bound	Upper Bound
Total	23.7 %	27.0 %
Refreshment	17.6 %	19.8 %
Panel	49.7 %	58.8 %

7. Implementation of the ERiK-Surveys 2024

This chapter presents the implementation of the ERiK-Surveys 2024 and includes survey design, contacting, field work and response rates. The contact materials are available online in the DJI-FDZ (https://surveys.dji.de/). An overview of the contacting and the survey designs for the different populations is depicted in Table 7-1.

7.1 Questionnaires

In preparation for the ERiK-Surveys 2024, the ERiK team revised the questionnaires to adjust them to the changing field of study and to improve convenience for respondents. The challenge was to balance the need to retain or add questions that generate relevant knowledge about the field of ECEC in Germany with the goal of keeping the questionnaires reasonably short to increase willingness to participate. Table 7-1 shows that a reduction in length was achieved in most surveys.

While the length of the director questionnaire increased by two pages (resulting in 27 pages in total), we reduced the number of questions from 94 to 81. The questionnaires for pedagogical staff show a similar pattern: compared to the ERiK-Survey 2022 (16 pages) we added three pages but reduced the number of questions from 54 to 50. Family day-care workers received the largest reduction in length. Here we decreased the number of pages from 26 to 22 and the number of questions from 81 to 70 compared to 2022. The number of pages for youth offices remained at 18 as in 2022, but we reduced the number of questions from 57 to 49. The questionnaire for providers shows an inverse pattern: the number of questions remained at 55 but we were able to reduce the number of pages from 21 to 19.

As the last row of Table 7-1 shows, the reductions in questionnaire length also decreased participation times. Participation times for directors were slightly shorter compared to the ERiK-Surveys 2022. Processing times for pedagogical staff were slightly longer than in 2022, which makes sense considering the questionnaire was three pages longer. By contrast, the median completion time for family day-care workers and youth offices was over ten minutes shorter than in the ERiK-Surveys 2022. A similar trend can be observed in the sample of providers.

7.2 Field Period

Overall, the field period for the five populations was from March to August 2024 and thus longer and later than in the ERiK-Surveys 2022. The later start was mainly due to the scheduling of the tendering and award process for selecting the external survey institute. Furthermore, the objective was to synchronise the field period as closely as possible with the start of the Child and Youth Welfare Statistics survey (KJH statistics). The use of extended field phases was deliberate, aimed at increasing response rates by giving potential respondents more time and incorporating additional field measures.

Before the start of the actual field period, the ERiK team implemented different measures to raise awareness about the upcoming ERiK-Surveys 2024. These included advertising the surveys on the ERiK homepage and the DJI twitter/X channel, asking relevant stakeholders in the Working Group

Table 7-1: **Contacting and Survey Design of the ERiK-Surveys 2022 & 2024**

	Directors		Pedagogical Staff		Family Day-Care Workers		Youth Welfare Offices		Providers	
	2022	2024	2022	2024	2022	2024	2022	2024	2022	2024
Field period	February – April	March – July	February – April	March – July	January – March	April/May – August	January – March	April/May – August	January – March	March – July
No. of pages paper questionnaire	25	27	16	19	26	22	18	18	21	19
No. of questions	94	81	54	50	81	70	57	49	55	55
Advertisement	X	X	X	X	X	X	X	X	X	X
Reminder by post	X	X	X	indirectly	X	indirectly	X	X	X	X
Reminder by email	–	X	–	indirectly	X	indirectly	X	X	X	X
Telephone contact/ reminder	X	X	–	–	–	–	X	X	X	X
Incentives	X	X	X	X	X	X	–	–	–	–
Non-Response Survey	X	X	–	–	–	–	–	X	X	X
Mean number of days between first contact and participation (median)	28.4 (26)	42.1 (28)	26.5 (20)	33.5 (28)	32.5 (32)	35.8 (32)	39.8 (38)	44.8 (41)	29.8 (30)	39.2 (36)
Mean (online) processing time in minutes (median)	46.5 (43)	42.2 (39)	31.5 (28)	34.6 (31)	53.6 (42)	32.9 (30)	62.5 (40)	35.7 (27)	47.8 (32)	28.1 (25)

Note: Only complete cases were used to determine the processing time of the online questionnaires. Processing times exceeding 500 minutes were excluded from the calculations.

on Early Childhood Education ("AG Frühe Bildung") to further promote the ERiK-Surveys and advertising them in various trade journals like "KiTa aktuell", "Kindergarten heute", "Kleinstkinder" and "didacta" which are mostly aimed at directors, pedagogical staff and family day-care workers. Furthermore, we asked the BMFSFJ to send emails to youth offices and providers to promote participation in the ERiK-Surveys 2024.

To encourage participation, pre- and post-incentives were offered in some populations. Directors and pedagogical staff of day-care centres received seed packets with their invitation letters (see Table 7.2-1) as a pre-incentive. Family day-care workers were informed in their invitation letter about a post-incentive: upon completing the questionnaire, they could choose either a €10 BestChoice voucher or donate €10 to an institution of their choice. About one quarter selected the voucher, while the remainder opted for a donation, most frequently to Deutsche Kinderkrebshilfe (German Children's Cancer Aid).

Directors & Pedagogical Staff

As in the ERiK-Surveys 2022, no tranche concept was implemented. However, for logistical reasons, the invitation letters were sent out on two dates: 3,383 on 15 March 2024 and 14,290 on 19 March 2024. Figure 7.2-1 shows the cumulative number of responses from directors and pedagogical staff respectively. Pedagogical staff were contacted exclusively through directors, who were asked to forward the questionnaires, QR-Codes, and reminders to their pedagogical staff. Therefore, response rates grew at a similar pace in the two populations. After the initial invitation letter, a written reminder was mailed to all centres with missing responses (director and/or pedagogical staff) on 8 April 2024. Additionally, an email reminder was sent to all centres on 11 April 2024. A second and third reminder were sent by email on 10 May and 17 July 2024, respectively. With each reminder, directors were asked to also remind their pedagogical staff to participate in the survey. Additionally, centres were contacted by telephone from 4 June 2024 onwards to encourage participation and conduct nonresponse interviews in cases of refusal to participate (see Section 7.4).

The first online interview was completed three days after the invitation letters were sent out. The median number of days until response was 28 for both directors and pedagogical staff; for directors, this is very similar to the median response time in 2022, while pedagogical staff took longer to respond in 2024 than 2022 (see Table 7-1). Note that that the field times in 2024 were longer than in 2022. Responses for both populations rose steadily and reached a plateau about a month later. Noticeable increases followed the first reminder and again after the second reminder, when a spike occurred likely due to the processing of paper questionnaires. Subsequently, response rates flattened, with telephone reminders showing no noticeable effect. Towards the end of the field period, a slight increase was observed after the third reminder.

Youth Welfare Offices & Family Day-Care Workers

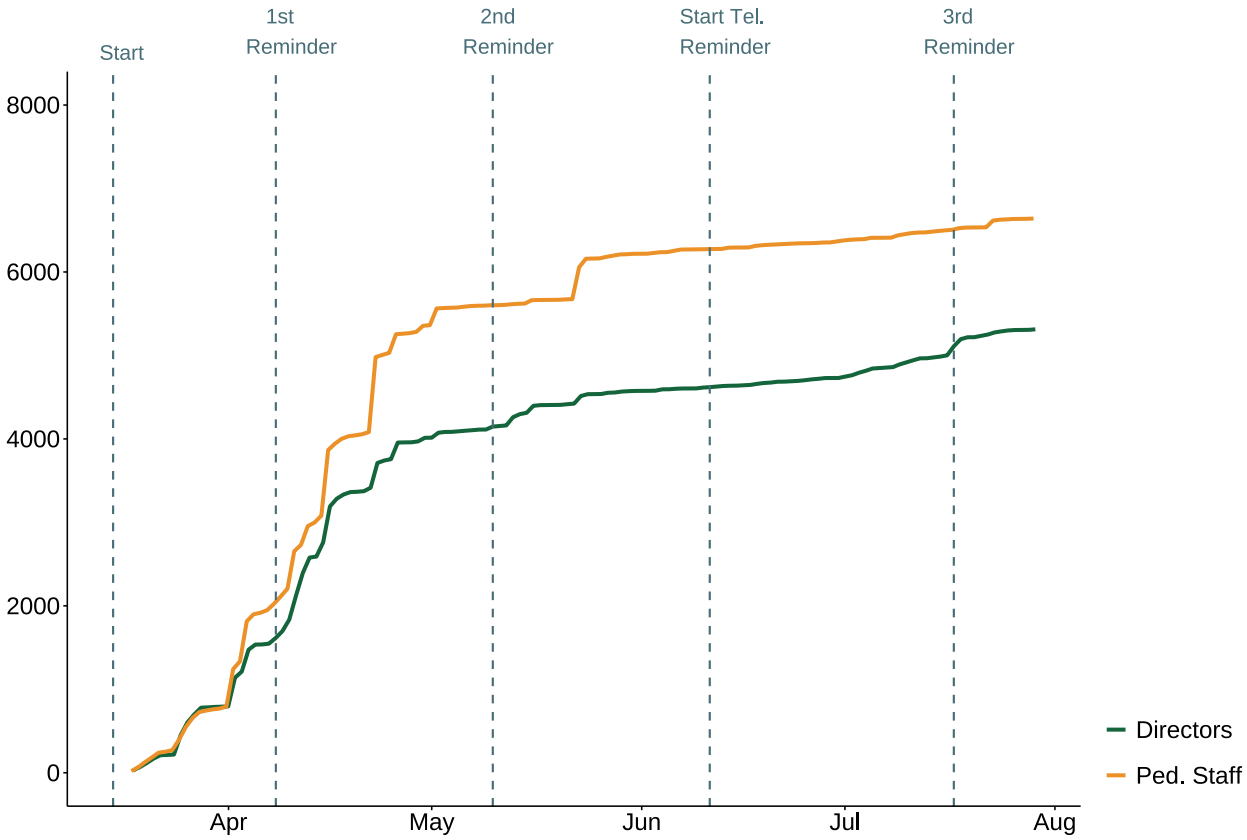
As with directors and pedagogical staff, no tranche concept was implemented for youth welfare offices. Instead, invitation letters were sent iteratively following preliminary telephone calls with the youth offices, with the majority of letters sent out on 22 April 2024, smaller batches sent on 24 April 2024 and 6 May 2024, and the last letter sent on 13 May 2024. The first online questionnaires were completed two days after dispatch; overall, youth offices responded after a median of 41 days while the median response time for family day-care workers was 32 days. These response times are similar to response times in the ERiK-Surveys 2022 (see Table 7-1). The youth offices received three written reminders to participate and encourage day-care workers to participate (23 May 2024, 13 June 2024, and 22 July 2024) and an additional telephone reminder (24 June 2024 – 22 July 2024). In cases of refusal to participate, respondents were asked to take part in a brief nonresponse interview (see Section 7.4).

Table 7.2-1: **Contents of invitation letters for different populations**

Population	Contents of Invitation Letter						
	Cover letter	Data privacy notice	Flyer	Recommendation letters by…	QR-Code / URL to online questionnaire	Paper version of online questionnaire	Seed packet as pre-incentive
Directors	Yes	Yes	Yes	› BMFSFJ › Bundesarbeitsgemeinschaft der freien Wohlfahrtspflege	Yes	14,295 (80.9 %)	Yes
Pedagogical Staff	Yes	Yes	Yes	› BMFSFJ › Bundesarbeitsgemeinschaft der freien Wohlfahrtspflege	Yes	33,560 (81.1 %)	Yes
Youth Offices	Yes	Yes	Yes	› BMFSFJ › Kommunale Spitzenverbände	Yes	Yes	No
Day-Care Workers	Yes	Yes	Yes	› BMFSFJ › Bundesverband für Kindertagespflege e.V.	Yes	18,885 (45.8 %)	No
Providers	Yes	Yes	Yes	› BMFSFJ › Bundesarbeitsgemeinschaft der freien Wohlfahrtspflege	Yes	No	No

Note: Invitation letters for pedagogical staff were sent to the directors who were asked to forward them to their pedagogical staff. Similarly, invitation letters for family day-care workers were sent to the youth offices which were asked to forward them.

Figure 7.2-1: **Cumulative response numbers of directors and pedagogical staff in the ERiK-Surveys 2024**

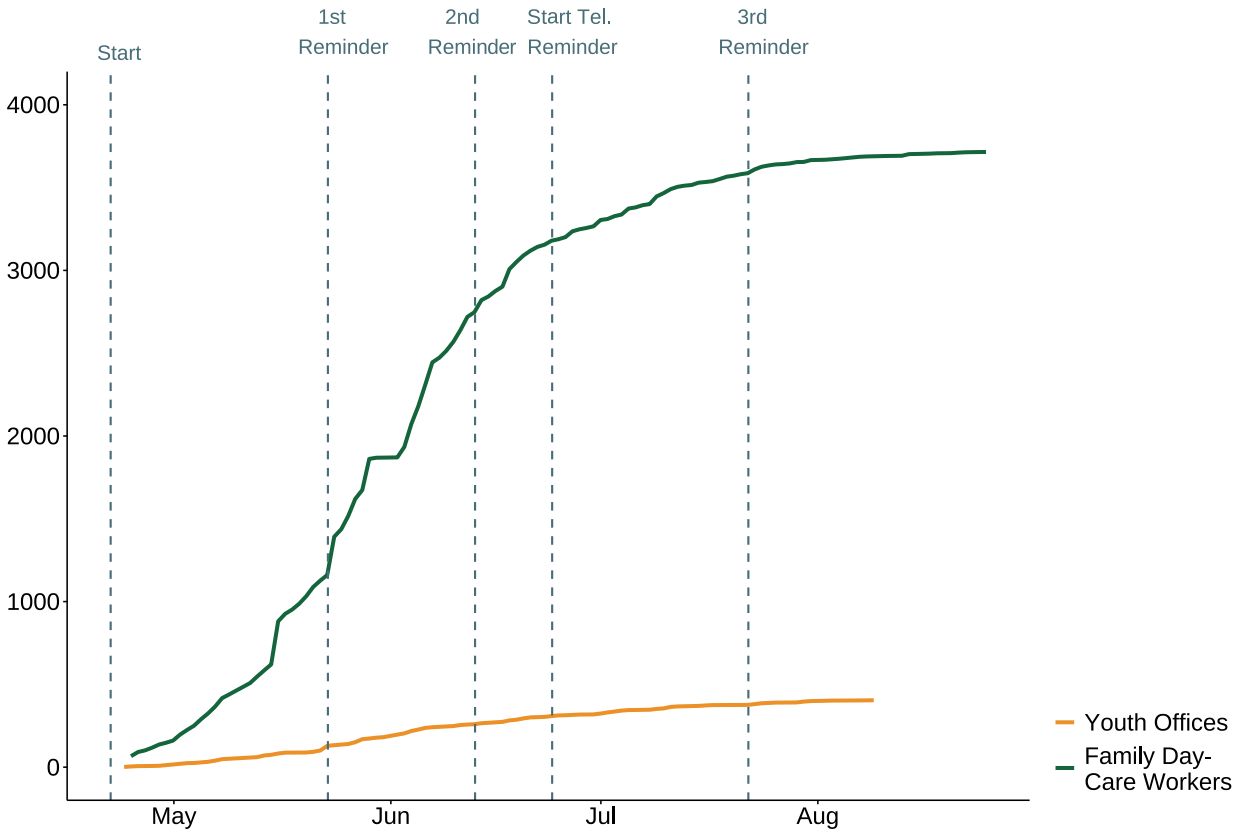


Most family day-care workers received the study materials indirectly via the youth welfare offices, which were asked to forward the pre-packed documents (see Table 7.2-1) to the day-care workers. In some cases, however, the materials were provided only by email and forwarded digitally. These procedures ensured that as many day-care workers as possible could be reached while reducing the burden on the youth offices as much as possible.

In the case of e-mail distribution, family day-care workers were able to register for the study via a direct link in the email, utilising a double opt-in procedure. This means they could register their e-mail address and would afterwards receive the link to the online questionnaire. Due to random duplicate checks by the survey institute, this distribution of individual access to the questionnaire was only conducted on a biweekly basis, which could have resulted in substantial waiting periods for some family day-care workers. It is likely that this procedure reduced the response rate in these cases compared to a process with instant automated access to the online questionnaire after registering one's email address, which is planned for the next ERiK-Surveys.

Figure 7.2-2 shows the cumulative response numbers of youth offices and family day-care workers. The youth office response curve rose steadily across the field period with no visible impact of reminders, plateauing roughly one month before the fieldwork ended. By contrast, family day-care workers showed marked increased after reminders, particularly the first, followed by smaller gains after the subsequent reminders and a final plateau from early August. The brief plateau in June may be due to temporary interruptions in data collection rather than participant behaviour.

Figure 7.2-2: **Cumulative response numbers of youth welfare offices and family day-care workers in the ERiK-Surveys 2024**

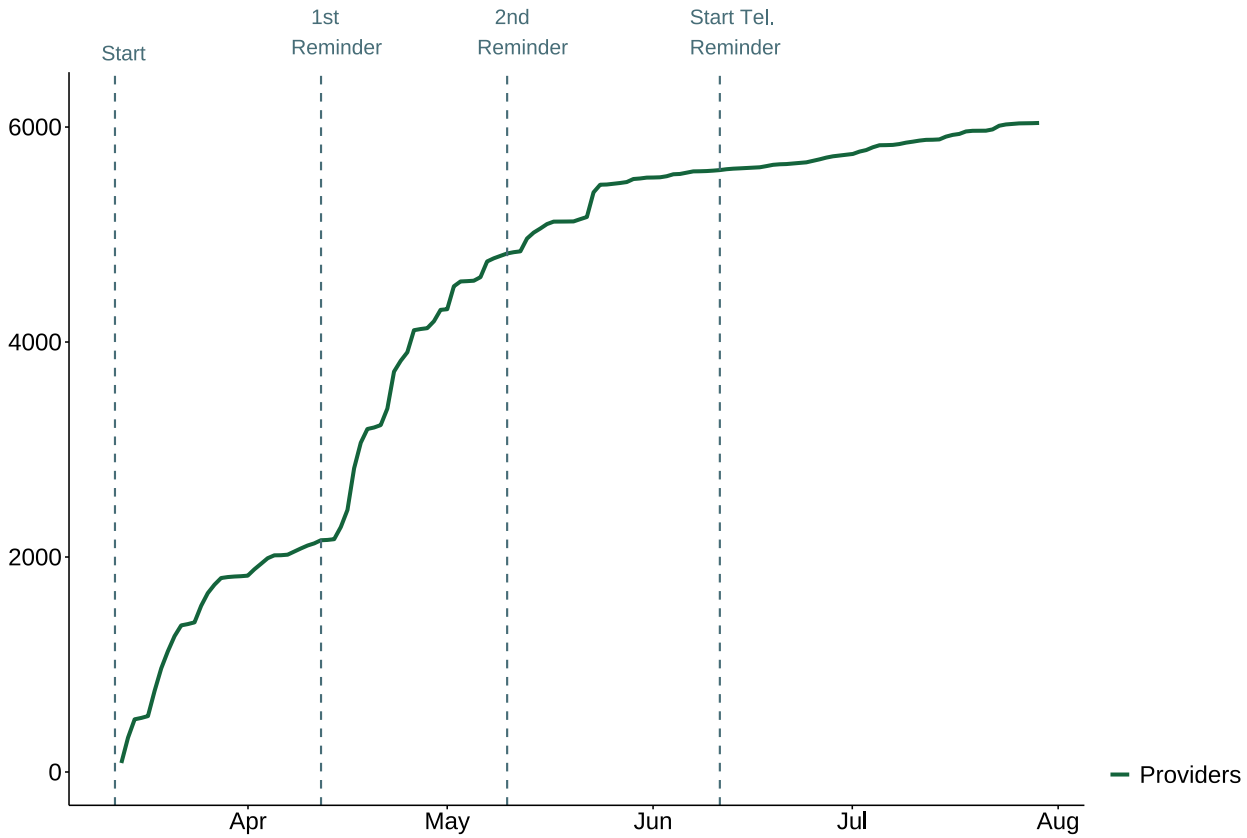


7.3 Providers

The first invitation letter was sent out to providers on 12 March 2024, leading to the first online questionnaire coming in one day later. The median response time of providers was 36 days – almost a week longer than in 2022 (see Table 7-1). However, the field times in 2024 were also longer for providers compared to 2022. In contrast to the other populations, there were only two written reminders in addition to the telephone reminder. Most providers received the first reminder via mail sent out on 12 April 2024 (N = 18,880). However, some were reminded via email on 17 April 2024 (N = 13,560). On 10 May 2024, the second reminder was sent out to a subset of providers (N = 10,960) via email. Telephone reminders took place from 11 June 2024 onwards. In cases where providers refused to participate, they were asked to take part in a brief nonresponse survey (see Section 7.4).

Figure 7.2-3 displays the cumulative response numbers of providers. The curve increases over time. The written reminders seem to have had an impact with response rates increasing steeply shortly after the first reminder. The lag between the reminder and the increase of the curve can be explained by the time it takes for the written reminders to reach the providers by mail. Overall, the response curve suggests that the written reminder strategy was effective, with participation increasing in clear waves following each contact attempt. Telephone reminders seem to have had less but still noticeable impact.

Figure 7.2-3: **Cumulative response numbers of providers in the ERiK-Surveys 2024**



7.4 Number of Completed Questionnaires and Response Rates

A total of approximately 5,300 directors, 6,600 pedagogical staff members, 3,700 family day-care workers, 400 youth welfare offices and 6,000 providers participated in the ERiK-Surveys 2024 (see Table 7.3-1). Of those about 4,900 directors, 6,500 pedagogical staff members, 3,600 family day-care workers, 380 youth offices and 5,900 providers completed the questionnaires in full according to our definition4, while the rest count as partial cases.

In comparison with the ERiK-Surveys 2020 and 2022, there has been an increase in the number of participants among directors and providers. The number of complete cases for youth offices increased compared to 2022 and is now back at the 2020 level, while the total number of cases (complete and partial) is between the numbers for 2022 and 2020. A decline in the number of complete cases was observed among pedagogical staff and family day-care workers in comparison to the ERiK-Surveys 2022.

4 According to the ERiK definition, complete cases are those that answered 80 % or more of the items and the relevant weighting variables in the questionnaire. If a respondent answered the weighting variables in full and 50-80 % of the items, it counts as a partially completed case. Less than 50 % of the items or an unanswered weighting variable leads to an incomplete/ break-off case.

Table 7.3-1: **Response Rates of the ERiK-Surveys 2024 compared to 2022 and 2020**

2024	Directors	Pedagogical Staff	Family Day-Care Workers	Youth Offices	Providers
Baden-Wuerttemberg	35 %	18 %	8 %	76 %	27 %
Bavaria	33 %	21 %	12 %	81 %	28 %
Berlin	**21 %**	10 %	5 %	**50 %**	21 %
Brandenburg	29 %	17 %	4 %	78 %	33 %
Bremen	26 %	10 %	**2 %**	**50 %**	23 %
Hamburg	**21 %**	**9 %**	5 %	**100 %**	20 %
Hesse	32 %	17 %	10 %	67 %	35 %
Mecklenburg-Western Pomerania	29 %	14 %	10 %	63 %	**19 %**
Lower Saxony	31 %	19 %	10 %	74 %	34 %
North Rhine-Westphalia	35 %	17 %	9 %	64 %	33 %
Rhineland-Palatinate	**43 %**	**23 %**	9 %	68 %	29 %
Saarland	31 %	15 %	6 %	**50 %**	37 %
Saxony	33 %	21 %	9 %	77 %	**46 %**
Saxony-Anhalt	29 %	16 %	8 %	**86 %**	38 %
Schleswig-Holstein	36 %	22 %	11 %	63 %	30 %
Thuringia	27 %	19 %	**13 %**	**86 %**	27 %
Overall	**30 %**	**16 %**	**9 %**	**71 %**	**30 %**
Total number of cases	**5,313**	**6,641**	**3,705**	**402**	**6,036**
2022					
Overall	**27 %**	**18 %**	**9 %**	**64 %**	**24 %**
Total number of cases	**4,832**	**7,116**	**3,927**	**366**	**5,166**
2020					
Overall	**30 %**	**19 %**	**20 %**	**83 %**	**16 %**
Total number of cases	**3,915**	**8,833**	**3,925**	**479**	**2,288**

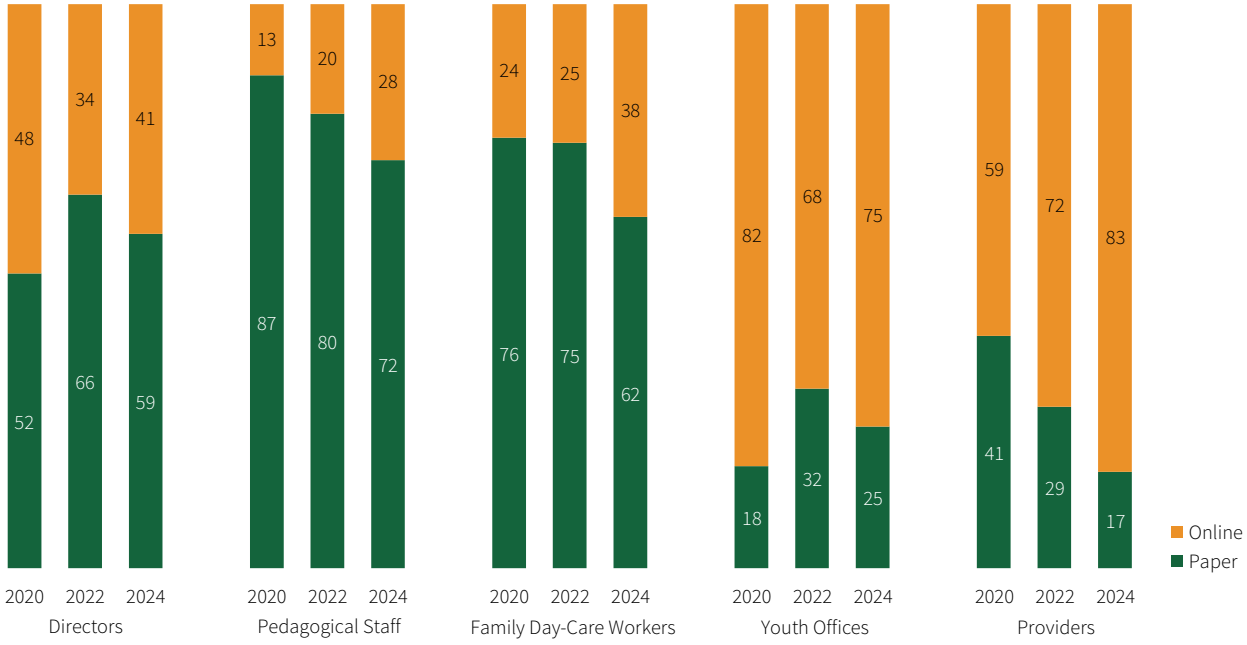
Note: Partial cases were counted here as having participated. Cases that are not eligible to participate were excluded from the calculation. Cases whose eligibility to participate is unknown were counted as eligible for the calculation. For the indirectly surveyed populations, the gross sample was reduced to those cases that had the opportunity to participate (to whom the survey was forwarded).

In comparison with the ERiK-Surveys 2022, the response rates (in accordance with the American Association for Public Opinion Research (AAPOR) definition number 2, with regard to the German context: Stadtmüller et al., 2019) observed an increase in most populations (see Table 7.3-1). Among directors, the response rate increased by 3 percentage points to 30 %, in comparison to the 2022 figures. Conversely, among pedagogical staff, there was a marginal decrease to 16 % (2022: 18 %). As illustrated in Table 7.3-1, there is considerable variation in response rates between federal states. For directors, they range from 21 % (Berlin and Hamburg) to 43 % (Rhineland-Palatinate), and for pedagogical staff, from 9 % (Hamburg) to 23 % (Rhineland-Palatinate). In the survey of family day-care workers, the response rate remained constant compared to the previous ERiK-Survey and continues to stand at 9 %. At the federal state level, response rates vary between 2 % (Bremen) and 13 % (Thuringia). The response rate from youth welfare offices exhibited a significant increase compared to the 2022 figures, now standing at 71 % (2022: 64 %). Furthermore, there are significant discrepancies between the federal states. The response rates vary from 50 % (Berlin, Bremen, Saarland) to 86 % (Saxony-Anhalt, Thuringia) and 100 % (the one youth office in Hamburg). The response rate among providers has increased significantly and now stands at 30 % (2022: 24 %). Again, considerable variations are evident among the federal states, with the range extending from 19 % in Mecklenburg-Western Pomerania to 46 % in Saxony, marking a disparity of up to 27 percentage points.

Three general trends are apparent from the ERiK-Surveys 2024. First, across the city-states of Berlin, Bremen and Hamburg response rates for the ERiK-Surveys 2024 are consistently lower than in other federal states, with this reduction being evident across all target populations.

The second apparent trend is that the mode, in which the respondents were invited to participate, influences the response rates (see e. g. Converse et al., 2008; Daikeler et al., 2020). As in the ERiK-Surveys 2022, the response rates for directors, pedagogical staff and family day-care workers were higher when they were allowed to choose between a paper and online questionnaire compared to only getting the online version (see also Section 3.4 in Herrmann et al., 2024). Consistently, the majority in these three populations also chose the paper questionnaire when they had the option (see Figure 7.3-1). In the ERiK-Surveys 2024, directors chose the paper questionnaire in 59 % of the cases, pedagogical staff in 72 % and family day-care workers in 62 % of cases. The youth offices and providers on the other hand preferred the online version even when they had the option of using the paper questionnaire. Here, only a small minority of these institutions chose the paper option (YWO: 25 % PRO: 17 %).

Figure 7.3-1: **Mode Choice by Population and Year in the ERiK-Surveys 2020, 2022 & 2024 (in percent)**



Note: The graph includes only cases that could choose between answering on paper or online and were either fully or partially completed; incomplete cases are excluded.

A third trend in the ERiK-Surveys seems to be a move towards more online participation across all populations. As Figure 7.3-1 shows, in three out of five groups the trend of increasing online participation started in 2020, in the other two it holds a least in comparison to 2022. For pedagogical staff, the share of online participation more than doubled from 13 % in 2020 (the lowest share across all years and populations) to 28 % in 2024. For providers, the online share rose from 59 % to 83 %, the highest proportion across all years and populations.

7.5 Nonresponse Surveys

Nonresponse interviews were conducted by telephone with one youth office, 513 directors, and 216 providers that had refused to take part in the survey, to investigate their reasons for not participating. The nonresponse surveys were integrated into the telephone reminder process: when individuals refused to participate, they were asked to answer questions about their reasons for not participating.

The single youth office that participated stated their current workload as well as participation in other surveys as reasons for their non-participation in the ERiK-Survey 2024. Among providers and directors of day-care centres, current workload (69 % and 74 % respectively) and time constraints for reasons other than work (42 % and 55 % respectively) were the most frequently mentioned reasons for not participating.

Notably, 12 % of both providers and directors named questionnaire complexity as a reason for non-participation, indicating that while the instrument was generally manageable, a minority of potential respondents perceived it as a barrier. Moreover, the proportion of providers stating that they had not received a questionnaire dropped from 36 % (2022) to 2 %, suggesting that the sampling frame (i.e. quality of the list of providers used for first contact) had improved compared to 2022. Taken together, the results of the nonresponse survey indicate that time constraints constitute the main obstacle across populations, mainly due to high workload.

8. Data Quality of the ERiK-Surveys 2024

8.1 Limited Comparability of the 2020, 2022 and 2024 Surveys

Over the three survey years of the ERiK project, there have been a few minor adjustments to the target populations. These were necessary, for example, to ensure a higher reachability of the population or to delineate the populations more sharply from one another. Such changes must be considered when interpreting the study results so that actual effects are not confused with artefacts due to altered target populations.

Changed Target Populations in 2022 for Directors & Pedagogical Staff

As part of the refinement of the ERiK project, an adjustment was made to the target population for pedagogical staff for the ERiK-Surveys 2022. Staff with management responsibilities were excluded from the survey in 2022, as their experiences and perspectives potentially differ from those of non-managerial pedagogical staff.

However, this change led to limitations in the comparability of the surveys from 2020 and 2022, as in the ERiK-Surveys 2020, pedagogical staff with management responsibilities were explicitly invited to participate in the staff survey. To nevertheless ensure comparability between the two survey years, cases of pedagogical staff with management responsibilities were subsequently excluded from the 2020 data (16 % of net cases). Furthermore, the design weights of pedagogical staff were adjusted for the ERiK-Surveys 2020 to correct the selection probabilities for pedagogical staff accordingly.

In addition, the adjusted definition of the target population for pedagogical staff from the ERiK-Surveys 2022 is not compatible with the corresponding definition in the Child and Youth Welfare Statistics (KJH-Statistik). This means that calibration with the marginal distributions of official data is not possible for the 2022 survey.

To rectify the lack of any calibration option, staff with management responsibilities were re-included in the survey for the ERiK-Surveys 2024. However, the 2024 questionnaire for pedagogical staff explicitly asked about management responsibilities so that they could be excluded from the analysis. Thus, data users can once again perform cross-sectional analyses for pedagogical staff with calibrated weights. However, in the present report, to ensure comparability with the 2020 and 2022 survey years, pedagogical staff with management responsibilities continue to be excluded from the analyses. Accordingly, uncalibrated weights are also used for the comparison and analysis of all three survey waves.

There were also slight adjustments to the target population for the director survey in the ERiK-Surveys 2022. In 2020, only directors who are primarily responsible were supposed to complete the questionnaire. In 2022, this definition was refined by using the „birthday method“ in cases of co-directorship tandems to ensure a randomized selection. This adjustment was maintained in 2024. The sociodemographic characteristics of these subgroups, which are relevant for calibration, were requested via an additional analysis of the KJH-Statistics to facilitate calibration. The calibrated weights for the director survey can therefore be used for cross-sectional analyses. Since a retrospective adjustment of the target population for 2020 is no longer possible, only the uncalibrated weights are used for the analyses in the present report.

Improved Sampling Frame for Providers of Childcare in 2022 and 2024

Compared to the provider survey in the ERiK-Surveys 2020, the use of lists provided by the German federal states in 2022 showed a clear improvement in the sampling frame. The better quality of the address data, which continued to improve for the ERiK-Surveys 2024, increases the precision of the estimates for providers compared to the 2020 survey year, but also means that comparisons with the reference year of 2020 should be made with caution. It is advisable to consider the results in light of the improved and updated database.

No Calibration for Providers in 2024

In 2020, no calibration of the weights was performed for the provider survey, as no official statistics on the distribution of providers were available. In 2022, calibration was first carried out based on the aforementioned address lists provided by the federal states. However, it is unclear to what extent these lists were complete and representative. Additionally, in the 2024 survey, we detected a potential problem with the use of provider size as a calibration variable. Due to an overly complicated design of the corresponding question in the questionnaire, some providers operating across federal states presumably did not report the number of their day-care centres in the respective federal state, but the nationwide total size of their provider organisation. This led to significant discrepancies between the size information reported by providers and the data available from the federal states. Due to these inconsistencies and lack of reliable provider statistics, calibration was omitted for the current provider survey. However, the impact of this change on the interpretation of the analyses is negligible.

Complete Survey of Family Day-Care Workers from 2022 Onwards

In the 2022 survey year, the survey of family day-care workers was changed from a sample survey (2020) to a survey of the entire population (complete survey), which was also maintained in 2024. The aim was to increase the response rate in this target population, which can only be contacted indirectly via the youth welfare offices and is thus hard to reach. Furthermore, the time and effort expended by the youth welfare offices in forwarding the surveys was to be reduced by eliminating the need for them to draw samples and by conducting the dispatch predominantly online in 2022. This methodological change should be considered when analysing and interpreting the results, as it may impact comparisons with data from the 2020 survey year.

8.2 Weighting

In order to generalise the findings from the ERiK-Surveys 2024, a weighting method is employed to adjust for variations in sampling and participation probabilities. As in the previous ERiK-Surveys, this process involves the integration of design weights to account for the complexities inherent in sampling procedures, nonresponse weights to address discrepancies in survey participation rates, and calibrated weights based on population distributions. For an overview, see Table 8.2-1.

Table 8.2-1: **Weighting Overview ERiK-Surveys 2024**

2024	Centres/ Directors	Pedagogical Staff	Family Day-Care Workers	Youth Offices	Providers
Design weight	Depends on the N of CEN in the federal state and the N of DIR in the CEN	Depends on the N of PST in the CEN, the n of sampled PST and the CEN DW	–	–	–
Nonresponse weight	Random forest with Probability Calibration utilising survey variables and a variety of regional variables at the district level				
Calibrated weight	For centres: 1. number of centres per type of provider (3 cat.) per federal state, 2. number of places for children (8 cat.)	–	1. number of FDW per federal state, 2. number of children in family day-care (6 cat.), 3. highest degree and hours of qualification course of FDW (8 cat.)	1. number of YWO per federal state, 2. Number of residents per YWO district (5 cat.), 3. FDW per YWO district (5 cat.), 4. day-care centres per YWO-district (5 cat.)	–
Trimming of final weights	Trimmed at 99 %				
Names of final weighting variables	CEN: nwe | DIR: nww	nww	nww	nww	nww

Design Weights

To adjust for the complex sampling designs of day-care centres, directors and their pedagogical staff (see Section 5.1) the ERiK team used design weights. Design weighting is carried out at the centre, director and pedagogical staff level.

The sampling probability for each centre is calculated as the quotient of the total number of centres in each federal state N_{CEN} and the number of centres in the sample n_{CEN} (see (1) in Table 8.2-2). The corresponding design weight is then the inverse of the selection probability (see (2) in Table 8.2-2).

To calculate the selection probability of directors, a quotient is formed by dividing the total number of directors per centre (N_{DIR}) by the number of directors who should participate (n_{DIR}), which usually corresponds to 1 unless several directors share equal responsibilities in one centre. The inverse of this quotient is then multiplied by the centre weight (see Table 8.2-2).

The sampling probability for pedagogical staff is calculated in a similar fashion and the quotient is formed from the total number of pedagogical staff per centre N_{PST} and the number of them who should participate in the survey n_{PST}. As with the director design weight, the inverse of the sampling probability is then multiplied by the centre design weight to generate the pedagogical staff design weight (see Table 8.2-2).

Table 8.2-2: **Selection Probability (1) & Design Weight (2) Calculations ERiK-Surveys 2024**

Centres	Directors	Pedagogical Staff
(1) $P(CEN) = \frac{N_{CEN}}{n_{CEN}}$	$P(DIR) = P(CEN) \cdot \frac{N_{DIR}}{n_{DIR}}$	$P(PST) = P(CEN) \cdot \frac{N_{PST}}{n_{PST}}$
(2) $DW(CEN) = \left(\frac{N_{CEN}}{n_{CEN}}\right)^{-1}$	$DW(DIR) = DW(CEN) \cdot \left(\frac{N_{DIR}}{n_{DIR}}\right)^{-1}$	$DW(PST) = DW(CEN) \cdot \left(\frac{N_{PST}}{n_{PST}}\right)^{-1}$

The surveys of providers, youth welfare offices and family day-care workers are designed as complete population surveys (in German: “Vollerhebung”). Therefore, no design weights are calculated for these populations.

Nonresponse Weights

For all target populations of the ERiK-Surveys 2024, the samples are drawn to give all individuals a known, non-zero probability of selection. However, the realized samples ultimately consist of voluntary participants, and survey participants typically differ from non-participants. Thus, to enable valid population inferences from the ERiK-Surveys 2024, the samples must be adjusted, for nonresponse bias. For this, propensity score adjustment is performed using inverse propensity weighted estimators (Ferri-García & Rueda, 2020). In the ERiK-Surveys 2024, „participation“ is defined as full completion of the questionnaire. By contrast, in 2022 for directors and family day-care workers, “participation” was defined as full or partial completion of the questionnaire.

Correcting for nonresponse bias involves formulating a model for each target population to estimate the probability of response. Ideally, a unique model would be specified for each outcome variable (Ferri-García et al., 2022), but this is unfeasible in a multi-topic survey like ERiK. Therefore, a single, comprehensive model is fitted for each population, incorporating all covariates that could be important for any outcome variable.

Traditional logistic regression models are prone to bias from model misspecification, such as violated linearity assumptions or unaccounted interactions, leading to underfitting (Gelein et al., 2018; Hastie et al., 2009). Instead, the ERiK-Surveys 2024 employ Random Forests (Breiman, 2001) to estimate response probabilities. This non-parametric, machine learning method constructs an ensemble of decision trees, introducing randomness via bootstrap sampling and by selecting random subsets of covariates for splitting. Random Forests can handle complex structures among a large number of covariates, reducing the risk of underfitting, though careful tuning is required to avoid overfitting and variance overestimation (Kern et al., 2019).

For the 2024 surveys, the sampling frames for all target populations were enriched with 150 district- and municipality-level administrative covariates (Bertelsmann Stiftung, 2025; Bundesinstitut für Bau-, Stadt- und Raumforschung [BBSR Bonn], 2024). As no direct sampling frame exists for family day-care workers, the model predicts the probability of a youth office forwarding the questionnaire to the workers in its district; the inverse of these probabilities serves as the nonresponse weight for this population. The complete sets of covariates are detailed in Table 11-2 in the Appendix.

A critical step in applying Random Forests is the configuration of their hyperparameters, which directly control the model’s behaviour and significantly impact its performance. For the ERiK-Surveys 2024, we advanced our methodology by replacing the random search used in 2022 with a more efficient Bayesian Optimization approach. Bayesian Optimization addresses the challenge of hyperparameter tuning by building a probabilistic model, typically a Gaussian Process (GP), that maps hyperparameters to the performance metric (in our case: ROC-AUC). Based on this surrogate model, an acquisition function intelligently selects the most promising hyperparameter set to evaluate next, balancing exploration of uncertain regions and exploitation of known high-performing areas (Snoek et al., 2012; Wu et al., 2019). This process iteratively updates the GP posterior, requiring far fewer evaluations than methods such as random search, which was used for the ERiK-Surveys 2022, to find optimal configurations.

Table 8.2-3: **Random Forests Hyperparameters**

Name	Description	Values
n_estimators	number of trees in the forest	[200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000]
min_samples_split	minimum number of samples required to split an internal node	[2, 5, 10]
min_samples_leaf	minimum number of samples required to be at a leaf node	[1, 2, 4]
max_features	number of features to consider when looking for the best split	[‚sqrt', None]
max_depth	maximum depth of the tree	[10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110]
bootstrap	Whether bootstrap samples are used when building trees	[True, False]

We conducted Bayesian Optimization over 20 iterations for each target population. We defined a grid with 6,480 hyperparameter combinations shown in Table 8.2-3. To ensure a precise and robust estimate of the model's generalisation error, we employed 20-fold cross-validation (CV). This higher number of folds, compared to the threefold CV used in 2022, allows for a more stable calculation of the mean cross-validation ROC-AUC (CV AUC), which served as the objective function for the optimizer. This means that 20x20 Random Forests were fitted with different hyperparameter combinations for each of the target populations of the ERiK-Surveys 2024. The hyperparameter sets yielding the highest CV AUCs were selected for each target population as the final configurations and are shown in Table 8.2-4.

Table 8.2-4: **Selected Hyperparameters of Random Forests**

	n_estimator	min_samples_split	min_samples_leaf	max_features	max_depth	bootstrap
Day-Care Centres	1200	2	4	sqrt	20	True
Pedagogical Staff	3000	2	1	sqrt	70	True
Family Day-Care Workers	2600	5	2	None	10	True
Youth Offices	1600	10	4	sqrt	110	False
Providers	200	10	4	None	10	True

Another methodological advancement for the ERiK-Surveys 2024 was the application of probability calibration to the Random Forest outputs (Zadrozny & Elkan, 2001). This post-processing step aims to produce more accurate and reliable estimates of the participation probabilities, thereby improving the quality of the nonresponse weights (Niculescu-Mizil & Caruana, 2005). The calibration process also reduces the variance of the predicted probabilities (see Table 8.2-5 columns 6 and 9). In the context of nonresponse weighting, higher variance of probabilities results in more extreme weights. Therefore, we argue that for the ERiK-Surveys 2024, where actual predictability of participation is somewhat limited even after careful hyperparameter tuning (see CV AUC in Table 8.2-5), this reduction of variance is desirable.

To evaluate the quality of the resulting Random Forest models, we compared their performance against logistic regression models. The logistic regression benchmarks were specified without regional-level variables, using only the core structural covariates (e.g. provider type, federal state). Table 8.2-5 presents the ROC-AUC for the optimized Random Forests as well as for the logistic regression models obtained when the final models are applied to the entire sample. For the Random Forests, the table also shows the mean CV AUC across all 20 cross-validation folds obtained during hy-

perparameter tuning. The results demonstrate that the Random Forests, tuned via Bayesian Optimization, clearly outperform the logistic regression benchmarks.

Table 8.2-5: **Results of Random Forest vs. Logistic Regression**

	Sample Size	Random Forest				Logistic Regression		
		Features	AUC	CV AUC	Standard Deviation	Features	AUC	Standard Deviation
Day-Care Centres	17560	178	0.84	0.63	0.09	30	0.66	0.14
Pedagogical Staff	41396	176	0.90	0.77	0.02	28	0.63	0.06
Family Day-Care Workers	567	170	0.96	0.67	0.07	22	0.54	0.10
Youth Offices	567	170	0.99	0.58	0.07	22	0.64	0.11
Providers	20152	174	0.79	0.64	0.02	26	0.61	0.08

Note: For categorical variables in the sampling frames, dummy variables were created for each item category.

Calibrated Weights

In a third stage, some of the nonresponse weights were calibrated with their respective population distributions to reduce variance and potentially address any remaining nonresponse or coverage errors. To do so, the ERiK team used iterative proportional fitting, or raking, (Deming & Stephan, 1940; Zaloznik, 2011) and population data from the 2023 KJH statistics (Destatis, 2023). As mentioned in Section 8.1, calibration weights for the ERiK Research Report V were only created for centres, family day-care workers and youth welfare offices.[5] The following characteristics were used for calibration:

- Centres: number of centres per type of provider per federal state, number of places for children
- Family day-care workers (FDW): number of FDW per federal state, number of children in family day-care, qualifications of FDW (highest degree and hours of qualification course)
- Youth welfare offices: Number of youth welfare offices per federal state, number of inhabitants, number of FDW, number of centres (in each case per youth office district)

The final weights (nww), either the calibration or nonresponse weights depending on the population, are then trimmed at the 99th percentile and normalised to the population size. The final weights are available to users in both trimmed and untrimmed forms.

8.3 Informative Value at Federal and State Level

The weighted data from the ERiK-Surveys 2024 provide a robust foundation for drawing generalisable conclusions about the target populations for the entire Federal Republic of Germany. Analyses at the national level are not subject to limitations.

However, state-specific evaluations for individual federal states (in German: Bundesländer) can be limited. The primary challenge in these sub-national analyses is achieving sufficient sample sizes, as small samples are inherently associated with a larger average sampling error (Groves & Lyberg, 2010). This can impair the statistical reliability, precision, and generalisability of the results. Small case numbers are particularly unavoidable in smaller federal states, where the total population of certain institutions (i.e. youth welfare offices or providers) is naturally low.

To ensure consistent and transparent handling of these limitations, the ERiK project adheres to a set of rules originally developed for the ERiK-Surveys 2020 (see Chapter 4.3 in D. D. Schacht et al.,

5 The ERiK team also produced calibration weights for directors and pedagogical staff but since they are not used in the ERiK Research Report V, their construction will not be described in this report.

2022) and simplified in 2022 (see Chapter 4.4 in Herrmann et al., 2024). These rules, detailed in Table 8.3-1, guide data analysis and reporting by balancing statistical accuracy, the fulfilment of distributional assumptions, and the preservation of data anonymity.

Table 8.3-1: **Rules for dealing with small case numbers**

Survey	Major limitations	Minor limitations
Directors	(n ≤ 75)	(n/N < 2 %)
Pedagogical Staff	(n ≤ 75)	(n/N < 0.5 %)
Family Day-Care Workers	(n ≤ 50) or (n/N < 5 %)	(n/N < 10 %)
Youth Offices	(n < 10) or (n/N ≤ 50 %)	–
Providers	(n < 50) or (n/N < 10 %)	–

Note: Major limitations: Comparisons across survey waves should not be made for this federal state. Minor limitations: Comparisons across survey waves for this several states should only be made if the there are no limitations in interpretability in the comparison year.

The landscape of analytical restrictions for the ERiK-Surveys 2024 remains largely consistent with that of 2022. The surveys of centre directors continue to have no analytical restrictions in any federal state.

For pedagogical staff, minor restrictions persist in the same states as in 2022 (Baden-Wuerttemberg, Bavaria, and North Rhine-Westphalia). These restrictions are not due to absolute case numbers but to a response rate that is low relative to the size of the target population in these large states. This is likely a consequence of the indirect contact method, where staff are contacted through their centre directors.

For family day-care workers, the pattern of minor and major restrictions has changed slightly. Minor restrictions are again present in Baden-Wuerttemberg Lower Saxony, North Rhine-Westphalia, and Saxony. In 2024, minor restrictions are also present in Hesse and Rhineland-Palatinate but not in Mecklenburg-Western Pomerania. Major restrictions, where the response rate was critically low or the absolute number of participants fell below a reliable threshold, continue to affect Berlin, Bremen, Saarland, Saxony-Anhalt, and Thuringia. Additionally, Brandenburg and Hamburg are also affected in 2024. This is attributed to the challenges of indirect sampling via youth welfare offices, which were even more pronounced in 2024.

Severe restrictions affect the survey of youth welfare offices in Berlin, Bremen, Hamburg, Mecklenburg-Western Pomerania, Saarland, and Saxony, which is a consequence of the small number of these institutions in these states. For childcare providers, severe restrictions due to very low absolute net case numbers (fewer than 50) persist in Bremen and Saarland.

In line with the approach established in 2022, federal states affected by minor limitations in interpretability can be evaluated and interpreted. However, data users should acknowledge the potential for self-selection bias in these samples. For states with major limitations in interpretability, the data are not sufficiently reliable for state-level inference.

For longitudinal analyses comparing an older survey (i.e. 2020 or 2022) with the 2024 survey, the following principles apply:

Comparisons should not be made for a federal state if there were major limitations in interpretability in either of the survey years used for comparison, or if minor restrictions were present in both the surveys used for comparison.

Comparisons are valid for states that have no limitations in both survey years used for comparison, or which have minor limitations only in one of the surveys.

For the ERiK research report V (Fackler et al., 2026), this means that no significance tests will be performed with regard to the differences between the ERiK surveys in 2024 and the ERiK surveys in 2022 and 2020 if there are significant restrictions for one of the federal states or minor restrictions

for both federal states. An overview of the informative value for all federal states across all surveys is provided in Figure 8.3-1.

Figure 8.3-1: **Informative Value of the ERiK-Surveys at the Federal and State Levels in 2024 & 2022**

2024	Directors	Pedagogical Staff	Family day-care workers	Youth offices	Providers
Baden-Wuerttemberg					
Bavaria					
Berlin					
Brandenburg					
Bremen					
Hamburg					
Hesse					
Mecklenburg-Western Pomerania					
Lower Saxony					
North Rhine-Westphalia					
Rhineland-Palatinate					
Saarland					
Saxony					
Saxony-Anhalt					
Schleswig-Holstein					
Thuringia					
Germany					

2022	Directors	Pedagogical Staff	Family day-care workers	Youth offices	Providers
Baden-Wuerttemberg					
Bavaria					
Berlin					
Brandenburg					
Bremen					
Hamburg					
Hesse					
Mecklenburg-Western Pomerania					
Lower Saxony					
North Rhine-Westphalia					
Rhineland-Palatinate					
Saarland					
Saxony					
Saxony-Anhalt					
Schleswig-Holstein					
Thuringia					
Germany					

No limitations Minor limitations Major limitations

9. Summary

The primary objective of this Methodological Report V was to provide a comprehensive description of the methodological framework for the ERiK-Surveys 2024. This encompasses the definition of the target populations, the construction and quality of the sampling frames, the sampling designs, and the subsequent concepts for data collection, evaluation, and weighting. By consolidating the information that was previously included in two separate methodological reports for the ERiK-Survey 2022 (see Gilg et al., 2023; Herrmann et al., 2024), this report offers a unified and in-depth reference for data users.

Significant aspects, including the refined construction of the official state lists, the updated nonresponse weighting using Random Forests with Bayesian Optimization, and the analysis of selectivity and response rates, have been carefully documented. Additional methodological details on the earlier ERiK-Surveys 2020 and 2022 remain available in the preceding methodological reports I to IV (Gilg et al., 2023; Herrmann et al., 2024; D. Schacht, 2022; D. D. Schacht et al., 2022).

The design of the ERiK-Surveys 2024 ensures that inference can be made from the data to make valid statements on the early education and care system in the whole of Germany. Readers of the forthcoming ERiK Research Report and users of the ERiK-Surveys 2024 data should consider the following key points:

- The ERiK Research Report V (Fackler et al., 2026) contains key findings that are generalisable to each target group. This is possible because of the data weighting processes that account for the complex sample design and nonresponse. It is recommended that all interpretations of analyses performed by any data user consider absolute case numbers and the dispersion of the data.
- Comparability across survey years (2020, 2022, 2024) is limited for some target populations due to deliberate methodological improvements and adjustments. These changes are detailed in this report (see Section 8.1) and must be considered when interpreting trends.
- The reliability of the ERiK-Surveys 2024 is limited for estimates in certain federal states and for specific surveys, primarily due to small sample sizes at the sub-national level. The rules for handling these limitations, consistent with the 2022 wave, are outlined to guide state-level analyses (see Section 8.3).
- The data for the 2024 survey wave will be made available through the DJI Research Data Centre (FDZ-DJI, www.surveys.dji.de) in due course. The survey data for 2020 and 2022 are already available there.

Overall, the ERiK-Surveys continue to provide a unique and high-quality data infrastructure. The 2024 wave, with its improved sampling frames and advanced weighting methodology, offers a robust and comprehensive foundation for a wide range of in-depth and multi-perspective analyses in the field of early childhood education and care, contributing significantly to evidence-based research and policy.

The synopsis provides a quick overview of the ERiK-Surveys 2024.

Synopsis ERiK-Surveys 2024	
ERiK Populations	Youth welfare offices (YWO), Family day-care workers (FDW), Providers (PRO) and Directors of & Pedagogical staff (PST) in Day-care centres (CEN)
Field time	› CEN/DIR, PST, PRO: March – July 2024 › YWO, FDW: April – August 2024
Target persons	› CEN/DIR: Person with the highest amount of managerial responsibilities in the centre › PST, FDW: selected person › YWO, PRO: Director, deputy director, one or more staff members
Survey institute	infas Institute for Applied Social Sciences
Survey instruments	Paper postal self-completion and access to the online version of the questionnaire (P&O) Access to the online version of the questionnaire (O)
Contacting	Target population-specific cover letter with information on study and data protection, postal reminder and telephone reminder (incl. nonresponse survey (for DIR, YWO, PRO))
Gross sample (GS); net sample (NC) and AAPOR 2 response rate[6] (RR)	› DIR: GS: 17,678; NC: 5,313, of which 4,931 complete (RR: 28 %) › PST: GS: 41,396; NC: 6,641, of which 6,475 complete (RR: 16 %) › FDW: GS: 41,235; NC: 3,705, of which 3,638 complete (RR: 9 %) › YWO: GS: 567; NC: 402, of which 380 complete (RR: 67 %) › PRO: GS: 20,665; NC: 6,036, of which 5,877 complete (RR: 29 %)
Weighting	› Design weighting for CEN/DIR, PST › Nonresponse weighting using random forest models for DIR, PST, FDW, YWO, PRO › Calibration/adjustment weighting for CEN, FDW, YWO
Data citation	› Total dataset (Gilg, Classe, Herrmann, Bopp, et al., 2025): Gilg, Jakob J./ Classe, Franz L./ Herrmann, Sonja/Bopp, Christine/Buchmann, Janette/Pachner, Theresia/Paulus, Andrea/Preuß, Melina/Romefort, Johanna/Selmayr, Anna/Tursun, Nadira/Kuger, Susanne (2025): ERiK-Surveys 2024. Dataset Version 1.0. Deutsches Jugendinstitut (DJI). [online at: https://www.doi.org/10.17621/erik2024_v01] › Survey specific datasets (see Bibliography): DIR: Gilg, Classe, Herrmann, Buchmann, et al., 2025, PST: Gilg, Classe, Herrmann, Pachner, et al., 2025, FDW: Gilg, Classe, Herrmann, Tursun, & Kuger, 2025, YWO: Gilg, Classe, Herrmann, Preuß, & Kuger, 2025, PRO: Gilg, Classe, Herrmann, Selmayr, et al., 2025

6 The AAPOR Response Rate Number 2 includes partial and complete cases (see Stadtmüller et al. (2019)).

10. Bibliography

Bertelsmann Stiftung. (2025). *Wegweiser Kommune 2024: Kommunale Daten für eine innovative Zukunft*. https://www.wegweiser-kommune.de/

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Buchmann, J., Ziesmann, T., & Drexl, D. (2022). HF-04 Stärkung der Leitung. In N. Klinkhammer, D. D. Schacht, C. Meiner-Teubner, S. Kuger, B. Kalicki, & B. Riedel (Eds.), *ERiK-Forschungsbericht II: Befunde des indikatorengestützten Monitorings zum KiQuTG* (pp. 109–123). wbv Media. https://doi.org/10.3278/9783763972999-07

Bundesinstitut für Bau-, Stadt- und Raumforschung. (2024). *INKAR – Indikatoren und Karten zur Raum- und Stadtentwicklung: Indikatorenübersicht*. https://www.inkar.de/documents/Uebersicht%20der%20Indikatoren.xlsx

Converse, P. D., Wolfe, E. W., Huang, X., & Oswald, F. L. (2008). Response Rates for Mixed-Mode Surveys Using Mail and E-mail/Web. *American Journal of Evaluation*, ***29***(1), 99–107. https://doi.org/10.1177/1098214007313228

Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. *Journal of Survey Statistics and Methodology*, *8*(3), 513–539. https://doi.org/10.1093/jssam/smz008

Deming, W. E., & Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, *11*(4), 427–444.

Dykema, J., Stevenson, J., Klein, L., Kim, Y., & Day, B. (2013). Effects of e-mailed versus mailed invitations and incentives on response rates, data quality, and costs in a web survey of university faculty. *Social Science Computer Review*, *31*(3), 359–370. https://doi.org/10.1177/0894439312465254

Fackler, S., Böwing-Schmalenbrock, M., Bopp, C., Classe, F. L., Meiner-Teubner, C., Kalicki, B., & Kuger, S. (Eds.). (2026). *ERiK-Forschungsbericht V: Befunde des indikatorengestützten Monitorings zum KiQuTG*. wbv Publikation.

Ferri-García, R., Beaumont, J.-F., Bosa, K., Charlebois, J., & Chu, K. (2022). Weight smoothing for nonprobability surveys. *TEST*, *31*(3), 619–643.

Ferri-García, R., & Rueda, M. d. M. (2020). Propensity score adjustment using machine learning classification algorithms to control selection bias in online surveys. *PloS One*, *15*(4), e0231500.

Gelein, B., Haziza, D., & Causeur, D. (2018). Propensity weighting for survey nonresponse through machine learning. In *13es Journées de méthodologie statistique de l'Insee (JMS).*

Gilg, J. J., Classe, F. L., Herrmann, S., Bopp, C., Buchmann, J., Pachner, T., Paulus, A., Preuß, M., Romefort, J., Selmayr, A., Tursun, N., & Kuger, S. (2025). *ERiK-Surveys 2024. German Youth Institute (DJI): Dataset Version 1.0.* https://www.doi.org/10.17621/erik2024_v01

Gilg, J. J., Classe, F. L., Herrmann, S., Buchmann, J., Paulus, A., & Kuger, S. (2025). *ERiK-Surveys 2024: Survey of directors. German Youth Institute (DJI): Dataset Version 1.0.* https://www.doi.org/10.17621/erik2024_l_v01

Gilg, J. J., Classe, F. L., Herrmann, S., Pachner, T., Paulus, A., Romefort, J., & Kuger, S. (2025). *ERiK-Surveys 2024: Survey of pedagogical staff. German Youth Institute (DJI): Dataset Version 1.0.* https://www.doi.org/10.17621/erik2024_p_v01

Gilg, J. J., Classe, F. L., Herrmann, S., Preuß, M., & Kuger, S. (2025). *ERiK-Surveys 2024: Survey of youth welfare offices. German Youth Institute (DJI): Dataset Version 1.0.* https://www.doi.org/10.17621/erik2024_j_v01

Gilg, J. J., Classe, F. L., Herrmann, S., Selmayr, A., Ulrich, L., & Kuger, S. (2025). *ERiK-Surveys 2024: Survey of providers of day-care. German Youth Institute (DJI): Dataset Version 1.0.* https://www.doi.org/10.17621/erik2024_t_v01

Gilg, J. J., Classe, F. L., Herrmann, S., Tursun, N., & Kuger, S. (2025). *ERiK-Surveys 2024: Survey of family day-care workers. German Youth Institute (DJI): Dataset Version 1.0.* https://www.doi.org/10.17621/erik2024_k_v01

Gilg, J. J., Gedon, B., Brusis, M., Buchmann, J., Drexl, D., Hoang, T., Jähnert, A., & Schacht, D. D. (2023). *ERiK-Methodological Report III: Conceptualisation and design of the ERiK-Surveys 2022.* wbv Publikation. https://doi.org/10.3278/9783763974566

Groves, R. M. (Ed.). (1989). *Survey errors and survey costs.* J. Wiley & Sons Inc.

Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present and future. *Public Opinion Quarterly*, *74*(5), 849–879.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.

Herrmann, S., Gilg, J. J., Classe, F. L., Schacht, D. D., & Kuger, S. (2024). *ERiK Methodological Report IV: Implementation and Data Quality of the ERiK-Surveys 2022.* wbv publication.

Kern, C., Klausch, T., & Kreuter, F. (2019). Tree-based Machine Learning Methods for Survey Research. *Survey Research Methods, 13*(1), 73–93.

Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. In S. Dzeroski (Ed.), *Proceedings of the 22nd international conference on Machine learning* (pp. 625–632). ACM Press. https://doi.org/10.1145/1102351.1102430

Olson, K., Smyth, J. D., & Wood, H. M. (2012). Does giving people their preferred survey mode actually increase survey participation rates? An experimental examination. *Public Opinion Quarterly*, *76*(4), 611–635. https://doi.org/10.1093/poq/nfs024

Schacht, D. (2022). *ERiK Methodological Report 1: Target Populations, Sampling Frame and Sampling Design of the ERiK-Surveys 2020* (1st ed.). wbv Media GmbH & Co. KG.

Schacht, D. D., Gedon, B., Gilg, J., & Kuger, S. (2022). *ERiK Methodological Report II: Implementation, Data Quality and Data Structure of the ERiK-Surveys 2020.* wbv Publikation. https://doi.org/10.3278/9783763973002

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. In F. Pereira, C. J. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 25: 26th Annual Conference on Neural Information Processing Systems 2012.* Curran.

Stadtmüller, S., Silber, H., Daikeler, J., Martin, S., Sand, M., Schmich, P., Schröder, J., Struminskaya, B., Weyandt, K. W., & Zabal, A. (2019). *Adaptation of the AAPOR final disposition codes for the german survey context.* GESIS – Leibniz-Institute for the Social Sciences. https://doi.org/10.15465/gesis-sg_en_026

Statistisches Bundesamt. (2023). *Statistiken der Kinder- und Jugendhilfe: Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2023.*

Wieschke, G. J., Lippert, K., & Kuger, S. (2025). *Methodischer Hintergrund der KiBS-Erhebung 2024: DJI-Kinderbetreuungsreport 2025.* Studie 8 von 8.

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., & Deng, S.-H. (2019). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. *Journal of Electronic Science and Technology*, *17*(1), 26–40.

Zadrozny, B., & Elkan, C. (2001). Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. *International Conference on Machine Learning*, *1*(5).

Zaloznik, M. (2011). *Iterative Proportional Fitting – Theoretical Synthesis and Practical Limitations* [Thesis (PhD)]. University of Liverpool, Liverpool. https://doi.org/10.13140/2.1.2480.9923

11. Appendix

Table 11-1: **Population and Samples of KiBS Surveys**

Target Population (0-6 years)	KiBS (Parent Survey)					
Survey-Year	2019 (W8)	2020 (W9)	2021 (W10)	2022 (W11)	2023 (W12)	2024 (W13)
Field Period	March – October	January – July	February– August	January – July	January – August	February – August
Population Size (0-6 years)	4.666.126	4.732.855	4.754.892	4.791.475	4.671.733	4.595.340
Sampling Frame	Indirect sampling through residents‘ registers of all German municipalities					
Sampling Strategy	Two-stage stratified cluster sampling					
Operational sample (gross, 0-6 years)	73.190	76.929	100.057	76.101	75.912	82.680
Net Sample (0-6 years)	19.981	20.847	18.790	19.772	18.898	21.559
Response Rate (0-6 years)	27,3 %	27,1 %	18,8 %	26,0 %	24,9 %	26,1 %

Table 11-2: **Completeness of contact data entries in address lists of providers and daycare centres**

	Centres						Providers					
	Email addresses			Telephone numbers			Email addresses			Telephone numbers		
	Original	2nd stage	3rd stage	Original	2nd stage	3rd stage	Original	2nd stage	3rd stage	Original	2nd stage	3rd stage
Baden-Wuerttemberg	0 %	70 %	75 %	96 %	96 %	97 %	72 %	80 %	83 %	97 %	98 %	98 %
Bavaria	0 %	0 %	27 %	98 %	98 %	99 %	0 %	39 %	45 %	82 %	93 %	94 %
Berlin	99 %	99 %	99 %	99 %	99 %	88 %	88 %	99 %	99 %	97 %	99 %	88 %
Brandenburg	95 %	95 %	95 %	98 %	99 %	99 %	99 %	100 %	100 %	99 %	99 %	99 %
Bremen	98 %	98 %	98 %	100 %	100 %	95 %	44 %	82 %	87 %	77 %	96 %	91 %
Hamburg	100 %	100 %	100 %	93 %	97 %	90 %	96 %	97 %	97 %	70 %	90 %	86 %
Hesse	87 %	87 %	89 %	96 %	97 %	98 %	84 %	88 %	89 %	99 %	99 %	100 %
Mecklenburg-West.Pom.	94 %	94 %	96 %	99 %	99 %	100 %	0 %	11 %	32 %	0 %	23 %	67 %
Lower Saxony	0 %	0 %	23 %	0 %	16 %	69 %	0 %	44 %	48 %	0 %	70 %	73 %
North Rhine-Westphalia	99 %	99 %	99 %	97 %	97 %	96 %	89 %	97 %	98 %	98 %	99 %	98 %
Rhineland-Palatinate	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	100 %	97 %	98 %	97 %
Saarland	98 %	98 %	98 %	98 %	98 %	99 %	66 %	76 %	81 %	66 %	89 %	93 %
Saxony	38 %	39 %	40 %	81 %	81 %	82 %	100 %	100 %	100 %	100 %	100 %	100 %
Saxony-Anhalt	92 %	92 %	93 %	99 %	99 %	99 %	75 %	86 %	87 %	75 %	77 %	89 %
Schleswig-Holstein	100 %	100 %	100 %	0 %	67 %	87 %	48 %	70 %	73 %	0 %	74 %	77 %
Thuringia	0 %	0 %	21 %	0 %	73 %	83 %	0 %	35 %	40 %	0 %	57 %	64 %
Germany	50 %	62 %	70 %	82 %	87 %	93 %	54 %	73 %	76 %	76 %	91 %	92 %

Table 11-3: **Overview of the weighting process for the ERiK-Surveys 2024**

	Centres / Directors	Pedagogical Staff	Family Day-Care Workers	Youth Welfare Offices	Providers
Design weight (DW)	multiplication of 1. total number of centres in a federal state divided by the number of contacted centres in the federal state by 2. the inverse proportion of contacted Ls and existing Ls in the centre.	multiplication of 1. the design weight of centres by 2. the inverse proportion of contacted Ps and the number of existing Ps in the centre (imputed if not available).	none (complete survey)	none (complete survey)	none (complete survey)
Nonresponse weight (NW)	Random Forest Model: Type of provider (8 cat.), municipality size, size of provider, federal state, survey mode, reminder by telephone, many regional variables.	Random Forest Model: Type of provider (8 cat.), municipality size, size of provider, federal state, survey mode, reminder by telephone, many regional variables.	Random Forest Model (Forwarding probability of YWOs): Federal State, municipality size, participation of YWO in 2023, many regional variables.	Random Forest Model: Federal State, municipality size, participation of YWO in 2023, many regional variables.	Random Forest Model: Federal state, Type of provider (8 cat.), size of provider/ number of centres, municipality size, many regional variables.
Calibrated weight (CW)	For centres: 1. type of provider (3 cat.) per federal state, 2. number of places for children (8 cat.).	none	1. number of FDW per federal state, 2. number of children in family day-care (6 cat.), 3. highest degree and hours of qualification course of FDW (8 cat.).	1. number of YWO per federal state, 2. number of residents per YWO-district (5 cat.), 3. FDW per YWO-district (5 cat.), 4. day-care centres per YWO-district (5 cat.).	none
Trimming	> 99 percent (< 1 percent if necessary)				

List of Figures

7.2-1 Cumulative response numbers of directors and pedagogical staff in the ERiK-Surveys 2024 27

7.2-2 Cumulative response numbers of youth welfare offices and family day-care workers in the ERiK-Surveys 2024 28

7.2-3 Cumulative response numbers of providers in the ERiK-Surveys 2024 29

7.3-1 Mode Choice by Population and Year in the ERiK-Surveys 2020, 2022 & 2024 (in percent) 31

8.3-1 Informative Value of the ERiK-Surveys at the Federal and State Levels in 2024 & 2022 40

List of Tables

4.1-1	Additional information in the lists of names and addresses of day-care centres and provider, by federal state	14
4.1-2	Sampling Frames ERiK-Survey 2024 and 2022: Day-Care Centres	14
5-1	Sampling and Survey Design ERiK-Surveys 2024	17
6-1	Upper and Lower Bounds of Response Rates	21
7-1	Contacting and Survey Design of the ERiK-Surveys 2022 & 2024	24
7.2-1	Contents of invitation letters for different populations	26
7.3-1	Response Rates of the ERiK-Surveys 2024 compared to 2022 and 2020	30
8.2-1	Weighting Overview ERiK-Surveys 2024	35
8.2-2	Selection Probability (1) & Design Weight (2) Calculations ERiK-Surveys 2024	35
8.2-3	Random Forests Hyperparameters	37
8.2-4	Selected Hyperparameters of Random Forests	37
8.2-5	Results of Random Forest vs. Logistic Regression	38
8.3-1	Rules for dealing with small case numbers	39
11-1	Population and Samples of KiBS Surveys	45
11-2	Completeness of contact data entries in address lists of providers and daycare centres	46
11-3	Overview of the weighting process for the ERiK-Surveys 2024	47

ERiK-Methodological Report V

Design and Implementation of the ERiK-Surveys 2024

The ERiK-Methodological Report V is the fifth in a series of methodological reports related to the ‘Entwicklung von Rahmenbedingungen in der Kindertagesbetreuung – indikatorengestützte Qualitätsbeobachtung (ERiK)’ study. The report contains all background information on the methodology of the ERiK-Surveys 2024.

Researching children, youth and families at the intersection of science, policy, and professional practice

The German Youth Institute (DJI) is one of the largest social science research institutes in Europe with an experience of over 60 years. The DJI conducts empirical studies into the life situations of children, young people and families, and provides policy advice to the German national government, the German federal states and local authorities as well as key impulses for professional practice.

Authors:
Franz Classe, Jakob J. Gilg, Caterina L. Sanchez Steinhagen, Sonja Herrmann, Susanne Kuger

wbv

ISBN: 978-3-7639-7996-7