COHORT '18
Growing Up in Hungary

HUNGARIAN DEMOGRAPHIC RESEARCH INSTITUTE

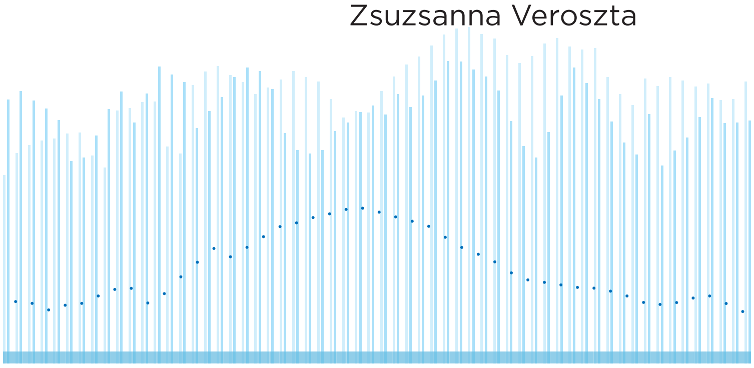
WORKING PAPERS

ON POPULATION, FAMILY AND WELFARE

Nº 41

GROWING UP IN HUNGARY
COHORT '18 HUNGARIAN BIRTH COHORT STUDY
TECHNICAL REPORT 3.
6-MONTH WAVE

Edited by
Zsuzsanna Veroszta



SZÉCHENYI 2020

HUNGARIAN GOVERNMENT

European Union
European Social Fund

INVESTING IN YOUR FUTURE

WORKING PAPERS

ON POPULATION, FAMILY AND WELFARE

Nº 41

GROWING UP IN HUNGARY
COHORT ’18 HUNGARIAN BIRTH COHORT STUDY
TECHNICAL REPORT 3.
6-MONTH WAVE

Edited by
Zsuzsanna Veroszta

Hungarian Demographic Research Institute
2023

ISSN 1588-3396
ISBN 978-963-959-770-9

Director: Zsuzsanna Makay
Series editor: Róbert Iván Gál
Reviewed by: Zsuzsanna Makay
Translation: Lukács Hayes
English proofreading: Clive Liddiard

Suggested citation:
Zsuzsanna Veroszta, Julianna Boros, Ildikó Fábián, Balázs Kapitány, Réka Kis, Krisztina Kopcsó,
Fruzsina Ökrös, Zsolt Spéder and Laura Szabó (2023):
Growing up in Hungary - Cohort ’18 Hungarian Birth Cohort Study. Technical report 3. 6-month wave. *Working Papers on Population, Family and Welfare, No. 41*
Hungarian Demographic Research Institute, Budapest.

Address: Hungarian Demographic Research Institute
Budapest, Buday László utca 1–3. 1024 Hungary
nki@demografia.hu
www.demografia.hu

Cohort ’18 Hungarian Birth Cohort Study
info@kohorsz18.hu
www.kohorsz18.hu

HUNGARIAN **DEMOGRAPHIC**
RESEARCH INSTITUTE

The research is financed within the framework of the EFOP 1.9.4. – VEKOP-16 invitation (renewing methodology and informatics in the social sector) issued by the Hungarian Ministry of Human Capacities.

The work of Fruzsina Ökrös was prepared with the professional support of the Doctoral Student Scholarship Programme of the Cooperative Doctoral Programme of the Ministry of Innovation and Technology, financed from the National Research, Development and Innovation Fund.

DOI: 10.21543/WP.2023.41

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. STRUCTURE OF THE 6-MONTH WAVE

1.1. POPULATION AND SAMPLE OF THE STUDY WAVE

The population of the Cohort '18 Growing Up In Hungary consists of children born in Hungary and their families in the period between 1 April 2018 and 30 April 2019. In each wave of the longitudinal study conducted by the Hungarian Central Statistical Office's Demographic Research Institute (HCSO HDRI), respondents from the initial sample are re-interviewed. The sampling frame of the study is Hungary's system of district health visitors. In addition to those respondents, the sample of the second study wave, conducted when the child was aged 6 months, was extended to include mothers from the core population of the pregnancy wave who had been unable to respond at the time of the first wave (e.g. due to premature birth, etc.). They were subsequently integrated into the research sample at the time of the 6-month survey by means of a so-called proxy questionnaire. Overall, the total population covered by the survey, including those respondents covered by the proxy questionnaire at the 6-month stage, consists of 8,717 mothers and 8,844 children.

The second wave of data collection in the longitudinal study (including the extension of the sample outlined in the previous paragraph) was based on a repeat survey of those women who had previously been interviewed while pregnant, this time 6 months after the birth. Any who had moved abroad since the pregnancy survey were not followed up and were thus excluded from the study sample. However, those who had moved within the country were kept in the sample, and considerable effort was made to follow them up successfully (see section 2), resulting in a very low drop-out rate. Overall, maximum size of the weighted database at 6-months consists of 8,311 primary caregivers and 8,424 children in their care. Naturally, this is a 'theoretical maximum': the useful number of items for each variable or research question is, in all cases, slightly lower (e.g. not all 'primary carers' are mothers, and so the number of items dealing with *mothers* is immediately slightly lower; in other cases, data from the prenatal wave are missing).

1.2. TIMING OF THE DATA-COLLECTION WAVE

Data collection for the second (6-month) wave of the Cohort '18 Study took place between June 2018 and November 2019. The survey, organized and administered by the health visitors, was timed to occur when the child was aged 6 months. The window for data collection ranged from 1 week before the child turned 6 months to 3 weeks after: a time frame of 1 month was therefore available for the survey. For premature infants, the enquiry period was adapted to the actual time of birth. The survey drew on the information received by the health visitor at the time of the birth of the child. As in the previous wave of the study, the organization and implementation of data collection was carried out by those health visitors involved in the study.

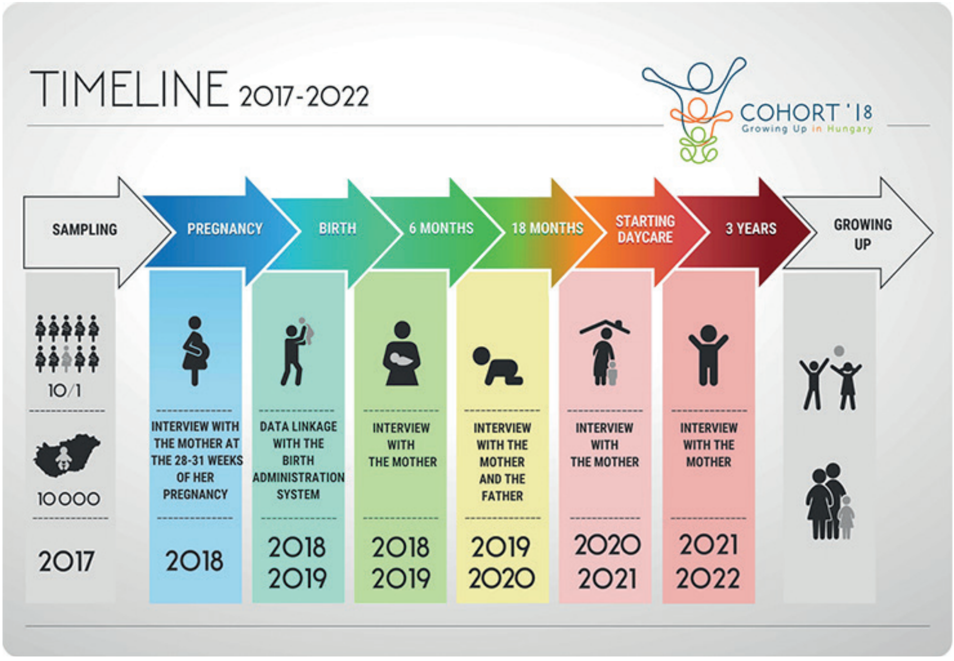


Figure 1
Waves of the Cohort '18 Study

1.3. QUESTIONNAIRES AND DATA-COLLECTION METHODS

The data-collection process was similar to that used in the previous, prenatal phase. Following contact initiated by the health visitor, the sampled mother or primary caregiver of the 6-month-old child was interviewed and the main questionnaire completed: generally, the mother's responses during the face-to-face interview were recorded on paper or online by the interviewer, and this was followed there and then by the respondent completing the self-administered questionnaire and by the health visitor filling in the corresponding health visitor's questionnaire. Personal contact data were recorded by the health visitors on index cards with identifiers. These were kept separate from the research data and were updated whenever there was a change of address, etc. The information was then passed on to HDRI.

While there was a core research questionnaire, different versions of the 6-month questionnaire were available for the birth mother and for any other primary carer. The self-administered questionnaire was completed only in the case of previous respondents. In the case of twins (triplets), the questionnaire items for the first, second or third child in the child module of the main questionnaire were applicable.

In addition to the main 6-month questionnaire, a short proxy questionnaire was used for any mothers who had been included in the initial population, but who for technical reasons had not been included in the pregnancy survey. The data content of this questionnaire was subsequently added to the pregnancy database.

New respondents who joined the study at the time of the 6-month wave – e.g. a primary caregiver who was not the biological mother or a mother who had not been included in the prenatal survey – had to complete a consent form before they could participate in the research (as in the previous phase). HDRI ensures that all personal data contained therein is kept separate from research responses. The linking of the data series and the identification is done using individual identification numbers (tokens).

Table 1

Questionnaires used in the 6-month wave

Name of the tool	Respondent	Conditions of application	Recording method	Data type
6-month main questionnaire of birth mother	Birth mother	The child is being raised by the biological mother	Paper-based oral interview with subsequent online recording or direct online recording	Research
6-month main questionnaire of other primary caregiver	Primary caregiver of the child	The child is being raised by someone other than the biological mother	Paper-based oral interview with subsequent online recording or direct online recording	Research
Self-administered questionnaire	Birth mother	The child is being raised by the biological mother	Paper-based, subsequently recorded by HDRI	Research
Prenatal-wave proxy questionnaire	Birth mother	No query during pregnancy	Paper-based oral interview with subsequent online recording or direct online recording	Research
Health visitor questionnaire	Health visitor	In all cases	Online recording	Research
Consent form	Birth mother or other primary caregiver joining the study sample	No enquiry during pregnancy or the child is not being raised by the biological mother	Paper-based, subsequently recorded by HDRI	Personal
Address card	Birth mother or other primary caregiver	In all cases	Paper-based oral interview with subsequent online recording or direct online recording	Personal

1.4. DATA PROTECTION

Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Regulation (EC) No 95/46/EC provides the legal framework for the Cohort '18 study regarding data protection. Its guidelines set out the requirements for the processing of any personal data for scientific research and statistical purposes that differ from the main rules. Given that the processing of personal data by the Cohort '18 study is deemed to be for scientific and statistical purposes, the study will apply the specific rules set out in that regulation.

The data controller of the Cohort '18 study is the Hungarian Central Statistical Office's Hungarian Demographic Research Institute, which guarantees an appropriate level of security in terms of the processing and storage of personal data, on the basis of the Data-Collection Protocol of the study. The main data-protection element of the research is anonymization through unique personal identifiers: this process combines the data collected in the different phases of the longitudinal research into a single research database with individual series, but without the inclusion of personal data. The processing of personally identifiable data – consent forms and their accompanying index cards – is kept separate throughout the study. Participants in the research are informed about the handling and security of their data in a Privacy Statement, which clearly sets out the legal framework, in particular with regard to data-protection safeguards. The legal basis for data processing is the consent form signed by all respondents, which allows the use and linkage of their personal data for research purposes throughout the study. The consent form also includes an index card for providing and updating contact details. For those respondents who are under the age of consent, there is a modified version of the consent form, to be signed by the respondent's guardian. Without a valid, signed consent form, the data collection is invalid and any data recorded will not be included in the study database.

The scientific ethical compliance of the Cohort '18 study is supported by a three-member Ethics Committee. The Ethics Committee is responsible for providing background support for the research, by commenting on and validating any ethical issues concerned with methodological decisions and data use in the research. On the initiative of HDRI, this committee has an ad hoc validation role in the design phase of the research and in ethical issues concerning the handling and use of data.

2. DATA COLLECTION

In order to test the clarity of the questions in the main and the self-administered questionnaires, the appropriate positioning of sensitive questions and the layout of the main sections of the questionnaire, including questions and groups of questions, a pilot survey was conducted prior to the live survey of the second wave of the Cohort '18 study. The pilot questionnaire also allowed us to observe which questions were harder or took longer to answer, and which questions respondents were unable or unwilling to answer. The selection of the health visitors participating in the pilot study was mostly random, but representativeness in terms of the type of settlement was an important criterion. The pre-testing was communicated to the health visitors in a letter of invitation sent by post, requesting them first to ask a mother with a 6-month-old child in their district to complete the questionnaire, and then to document their experiences and comments on the survey. Once the health visitor agreed to participate, she was given a further briefing on how the pilot survey would be conducted, what to look out for and what to record. The testing package prepared and mailed to the health visitors included a document summarizing the instructions for the pilot survey, a letter of invitation for the mother selected, the main questionnaire and the answer sheet booklet, two self-administered questionnaires and the necessary A5 envelopes, as well as a stamped addressed A4 envelope, a sheet for feedback on general problems, a playsuit for the baby (donated by Lidl Hungary) and a voucher worth 18,000 HUF donated by the online platform *Gyerünk, anyukám!* for its so-called 'regeneration gym'. The voucher was also provided as a gift for everyone who participated in the subsequent live survey, but only those mothers who participated in the pilot research received a playsuit.

Subsequently, the health visitors recruited eligible mothers in their district using a pre-prepared letter of invitation from HDRI. In the pilot research phase, 62 inquiries were made: of these, 48 were tested on the questionnaire for biological mothers and 14 on the questionnaire for non-biological mothers.

In parallel with the pilot interview, the health visitors immediately jotted down their observations regarding both the main and the self-administered questionnaire. As the self-administered questionnaire was usually completed by the mother herself (meaning that the professional did not actually see the answers), two copies of the questionnaire were provided, so that the health visitor could help the mother if she had difficulty in answering the questions. A respondent's comments made while completing the self-administered questionnaire could be noted by the health visitor on the printed version of the general issues feedback sheet.

After the pilot survey, we asked the health visitors to record their experiences as soon as possible using the online evaluation interface. The online spreadsheet was used to record any specific comments that had been noted down on the questionnaire as the interviewer worked her way through the questions; but it was also important for the health visitors to share their own comments and impressions of the questionnaires. The feedback was collected in separate databases according to each type of questionnaire (main questionnaire, self-administered, proxy, twin questionnaire). After recording the information, the health visitors were asked to return the interview materials (the completed questionnaires and the experience sheet) and two copies of the contract in a prepaid, addressed envelope to HDRI. The information collected and the feedback received were of enormous help in the subsequent follow-up and finalization of the questions and in the design of the questionnaires.

In addition to the experience gained during the pilot phase, we also built on the opinions of experts in various fields concerned to finalize the research questionnaires. During the 6-month study wave in April 2018, the following nine experts provided written feedback on the questionnaires:

Zsuzsa Blaskó, senior research fellow, HCSO HDRI
Ildikó Danis, senior research fellow, Semmelweis University Institute of Mental Health
Beáta Dávid, scientific advisor, Hungarian Academy of Sciences Centre for Social Sciences, Institute for Sociology
Éva Gárdos, senior technical advisor, HCSO
Kálmán Joubert, senior researcher, HCSO HDRI
Ferenc Kamarás, chief statistical advisor, HCSO
Péter Róbert, lead researcher, TÁRKI
Károlyné Tokaji, department head, HCSO
Dr György Velkey, director general, Bethesda Children's Hospital.

2.1. ORGANIZATION OF THE FIELDWORK

As in the first wave of the Cohort '18 study, the fieldwork was carried out using the network of health visitors. In order to organize the data collection and carry out the logistical tasks, there was continuous cooperation and contact between the staff of HDRI, the lead health visitors and the health visitors in the districts sampled.

The flowchart below shows the different steps in the study wave. The first element in the flowchart – the pilot research phase – was described in the previous section, while the data collection for the 6-month survey is described below.

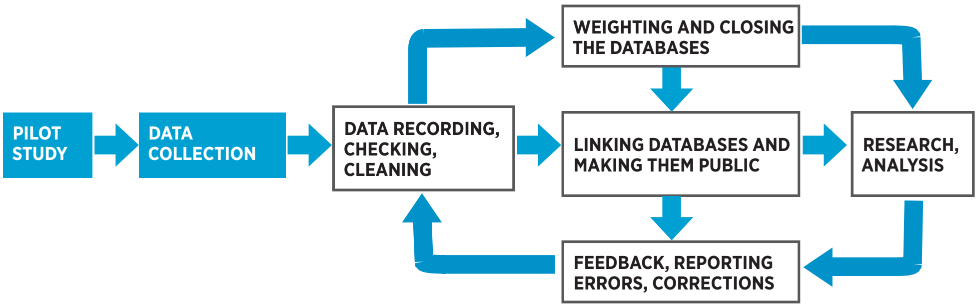


Figure 2
Flowchart of the Cohort '18 study, in sequence (steps 1 to 2)

Contacting the lead health visitors

All health visitors in the Hungarian system belong to a district, each of which is managed by a lead health visitor. Consequently, HDRI had to work closely and continuously with the head district health visitors, in order to support the flow of information and to organize their work, especially in terms of logistics. The most important element in this liaison was a regular management circular, which provided information on the progress of data collection: district heads were informed about the survey's data-collection schedule, deadlines and sampling rules. In addition, the circulars were used to organize training sessions for the health professionals.

Furthermore, as any health-visitor changes, district replacements and transfers were also linked to the district heads at the organizational level, the fieldwork staff paid special attention to cases where a new district head was appointed to a district that was participating in the research, or where a participant moved to a district where her old health visitor or district head was not responsible. In such cases, we tried to contact the participants as soon as possible and asked the new lead health visitor or health visitor to join the research and take on the data-collection tasks, either by email or by telephone.

Provision of logistical tasks

The delivery, distribution and collection of the survey packages was achieved using pre-established logistical points. Those logistical points set up during the prenatal wave – most of which were located in the district head's office or (especially in the case of geographically dispersed districts) at the district level – functioned in the same way during the 6-month data-collection wave (Szabó et al., 2021).

In each case, the research logistical tasks were communicated separately and directly to the relevant health visitors. As the quantity of material required for the 6-month wave was calculated on the basis of the number of women who had participated in the prenatal wave, we also had to reorganize the packages in the light of drop-outs and moves away: packages had to be reallocated from those settlements to which we had sent too many to those settlements that had received too few.

The packages for the 6-month wave of the Cohort '18 survey included the following:

A printed paper version of the main questionnaire, on which the interviewer could jot down the answers at the time of the interview for later online recording. The main questionnaire also included the proxy prenatal-wave questionnaire, to be completed if necessary, as well as the accompanying informed consent form.

A printed paper version of the self-administered questionnaire.

An envelope in which the completed self-administered questionnaire was sealed by the health visitor.

A letter of invitation congratulating participants on the birth of their child and informing them of the procedure for the 6-month survey.

A colourful information booklet briefly summarizing the aims of the study, the data-handling guidelines and the conditions of participation. In addition, colour illustrations were used to report some of the findings from the prenatal wave, based on the responses of the first 1,000 participants. We also informed respondents of the separate Cohort '18 telephone survey wave, in which fathers would be contacted when their child is aged 18 months; we also mentioned that this 18-month interview would be conducted not by the health visitor, but by a professional interviewer with a photo ID.

A hooded baby towel for the infant child – to be offered to the mother as a gift after completion of the survey (two in the case of twins).

A gift (tea bags) for the participants offered by NATURLAND Hungary Ltd.

A voucher that mothers could redeem for a postnatal health programme via the *Gyerünk, anyukám!* online platform.

Strengthening the commitment of health visitors and keeping in touch

One of the most important tools in coordinating the fieldwork and keeping the health visitors who collected the data up to speed was a regular newsletter with information about the tasks and deadlines related to the survey, as well as details of payments.

They could also contact our staff by email and telephone with any questions or problems that they might have encountered (e.g. questions about moving house, clarification of issues related to data recording), as part of a continuous 'helpdesk'. The private interface for health visitors created on the Cohort '18 study website was maintained, so that they could easily access the different questionnaires and browse the frequently asked questions section. In addition, in order to ensure continuous and smooth communication and the flow of information, the private Facebook group that was set up at the outset of the research (and accessible only to health visitors) continued to operate.

During the 6-month wave, the lead health visitors in charge of the logistical points and the health visitors continued to receive remuneration for their work. An increased fee was paid if the health visitor interviewed a mother of twins or if the mother interviewed had at most only 8 years of primary education. Furthermore, the health visitor was paid a higher rate if she undertook to visit and interview a mother who had moved outside the health visitor's district. In addition, any newly recruited health visitor into whose district a former pregnancy-wave participant had moved was also paid at a higher rate for interviewing her. Payments were scheduled quarterly by HDRI once a quarterly data-collection phase was complete, and were always settled on the basis of the number of questionnaires and consent forms received by the predetermined quarterly date.

In addition, we sent each health visitor who was participating in the research a Christmas card, and we have kept them all informed of the progress of the research and the partial results of the data collection on the official Facebook page and on the website of the Cohort '18 study.

During the 6-month wave, since the fieldwork was largely undertaken by the same health visitors, it was rare for any additional training of the health visitors to be needed: this was required only in isolated cases (e.g. where a new health visitor had been recruited) or in areas where, according to the feedback from the health visitor, the questioning was stalled for some reason among participating mothers. During the 6-month wave, between 12 June and 19 November 2019, research staff visited nearly 30 different sites, including Baja, Balatonfüred, Budapest III, VIII, X, XI, XIV and XXI districts, Diósd, Dorog, Ecser, Fehérgyarmat, Gödöllő, Győr, Hódmezővásárhely, Kaposvár, Komárom, Kunszentmárton, Miskolc, Orosháza, Paks, Pilisvörösvár, Siófok, Szeged, Székesfehérvár and Szentendre. The aim was to facilitate smooth fieldwork through refresher training. The training sessions took place in small groups, after prior consultation with the head district health visitors.

In 2018, in of the journal *Védőnő* (no. 28, a journal for health visitors), we published a one-page briefing on the launch of the 6-month wave; and in 2019 (vol. 29, no. 3), in addition to an article on the conclusion of the prenatal wave, we also reported on the development of the second wave of research. As well as those colleagues directly involved in the data collection, the articles allowed the whole network of health visitors to become better informed about the current state of the research.

Management of identifiers

During the 6-month data-collection wave, respondents – and, by default, health visitors – continued to participate using the tokens they had received when the data were collected during the pregnancy period. As already mentioned, these identifiers will remain unchanged for the duration of the study. We reported in detail on the criteria for assigning the identifiers and on the differences between tokens and serial numbers in the technical report on the prenatal-wave data collection (Szabó et al., 2021). It is important to note, however, that while the six-digit number used to identify a child during data collection in the prenatal wave was also used to identify a health visitor, certain factors (such as the mother moving away) meant that it was no longer appropriate for identifying a health visitor in the 6-month data collection: if the mother has moved to a district not covered by the study, and a new health visitor has had to be recruited, the identifier for the mother (or, in the case of twins, the children's identifiers) will have remained the same, but a new identifier will have been allocated to the health visitor.

There may also have been changes to the health visitor personnel within a district (e.g. due to death, retirement, change of career or relocation). In such cases, a newly arrived health visitor will have been given a four-digit identifier (the next in the sequence) that includes a number identifying the district.

2.2. SURVEYING PEOPLE WHO HAVE MOVED AND NEW ENTRANTS

In order to avoid attrition and sample attrition, both the health visitors and the fieldworkers made a special effort to follow up and maintain the sample of respondents who had moved. First of all, the folder compiled for the health visitors during the first wave of data collection contained the main details of the successful interviews conducted during the research, as well as details of planned future data-collection phases. Secondly, this folder was also used to collect index cards containing the respondents' contact and recruitment details. In the event of a move, these index cards, prepared in advance and supplied by HDRI, could be used by health visitors to record the new contact details of participants. Furthermore, a request was made to the health visitors in connection with any house move – which was also emphasized throughout during the online and face-to-face communication related to the study – that, as well as recording the change on the index card, the new contact details should also be entered on a dedicated online interface, allowing the address details in the HDRI database to be continuously updated, thus facilitating the success of future outreach to respondents.

In the period between the prenatal and the infancy (6-month) waves, health visitors reported a total of 660 respondents who had moved house. These included both instances where the person moved to a different settlement and cases where the person changed their address within the settlement where they lived. Of the 660 respondents who had moved house, 532 were successfully surveyed during the 6-month wave. Of those 532 surveys, 521 were completed with the biological mother (K1), and 11 with a primary carer who was not the biological mother (K2). Of the remaining 128, 28 mothers and two primary carers who were not the biological mother were successfully surveyed at the 18-month stage. In a further 29 cases, although it was not possible to interview them when their child was 18 months old, the mothers were still included in the initial database for the 3-year wave. In all, 61 people who had moved were permanently dropped from data collection during the 6-month wave; at the time of the 18-month wave, a further eight mothers were permanently dropped because the survey was unsuccessful.

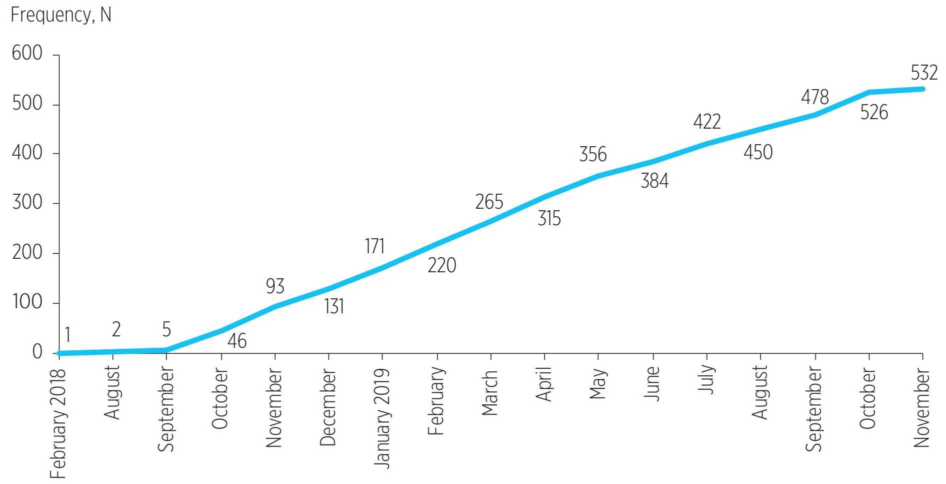


Figure 3

Number of successful interviews at the 6-month survey stage among mothers who had moved

During the prenatal-wave data collection, one major challenge was that some pregnant women who had attended the maternity services and had agreed to participate in the research did not have the opportunity to have their data collected before they gave birth. The main reasons for an interview not being conducted included premature childbirth, prenatal hospitalization, being abroad or even – in some instances – the illness of the health visitor. However, we still wanted to reach those mothers, and so we gave them the opportunity to join the study by completing an additional, condensed, retrospective (proxy) questionnaire when their child was 6 months old. Although those respondents who joined at the 6-month stage are not included in the prenatal cross-sectional analysis database, they are already full members of the study for the purposes of subsequent data-collection waves. It is important to note that if a new entrant joined the research at the 6-month stage, prior to data enrolment she was required to sign a consent form (which the other participants had completed at the outset of the prenatal data-collection wave), and so she was not required to sign a new consent form at the time of the 6-month wave. This was also the case if a child who was already enrolled had been handed over to a new carer in the first 6 months of life: if the new carer was contacted and confirmed her willingness to participate, she was also asked to sign the consent form. During the 6-month data collection, 394 proxy questionnaires were completed – 383 by the biological mother and 11 by a primary carer who was not the child's biological mother. It is important to emphasize that while 383 proxy questionnaires were completed by biological mothers, instances of twin pregnancies meant that the questionnaires actually refer to 388 children.

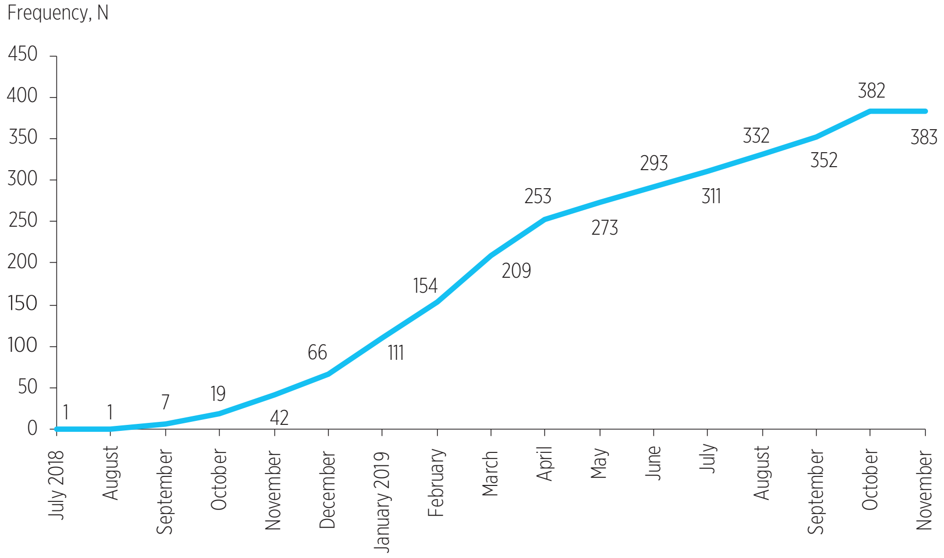


Figure 4

Number of proxy questionnaires completed by biological mothers

2.3. EVOLUTION OF SURVEYS

The drop-out rate

During the 6-month data-collection wave, there were 7,874 surveys carried out that had been preceded by successful contact during pregnancy. Of the 16 foreign-language contacts recorded at the pregnancy phase, five were continued during the infancy (6-month) wave. Although two further surveys were successfully carried out during the 6-month wave, these had to be excluded from the prenatal database due to lack of data from the first wave. Although these interviews form part of the 6-month cross-sectional database, they are not part of the prenatal database. To illustrate the drop-out, the five foreign-language interviews and the additional two interviews that were deleted from the first wave due to data shortages were removed from the 6-month database (n=7867).

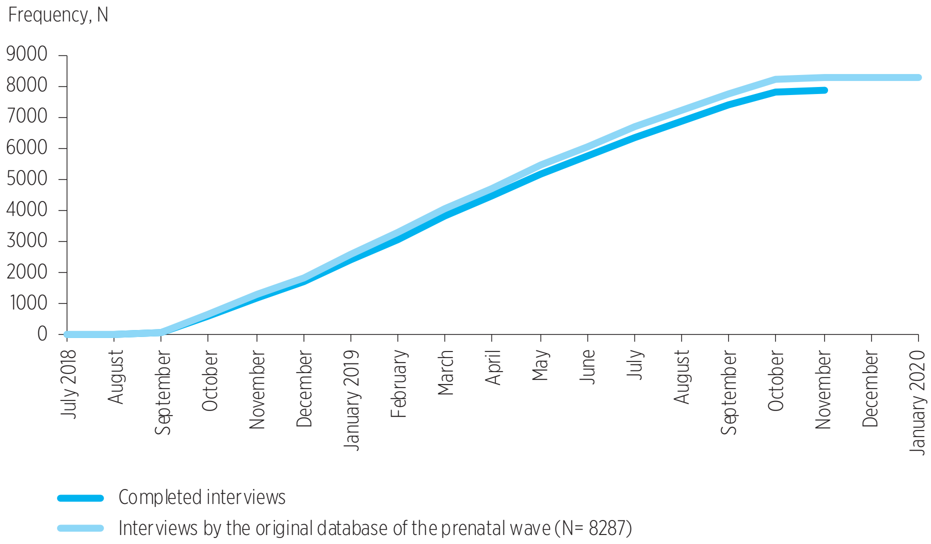


Figure 5

Evolution of the number of successful surveys

In all, then, there were 420 cases where contact was successfully made during the pregnancy wave but failed at the time of the 6-month wave. Among the drop-outs, there are two cases where contact was successfully made during the prenatal wave, but the mothers dropped out before the second wave of data collection, and so we do not know the actual date of birth of their child (and therefore cannot ascertain the drop-out date). These two cases are not included in this graph.

Table 2

Main reasons for early drop-out

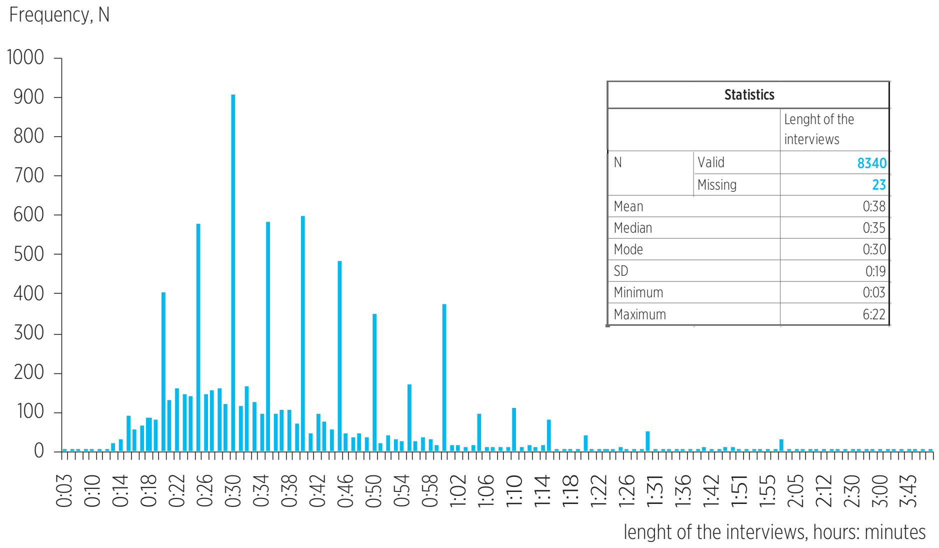
Reason for drop-out	N	%	Drop-out rate compared to the prenatal wave
Permanently dropped out of the longitudinal sample before the 6-month data-collection wave (e.g. late foetal death, stillbirth, conscious exit)	27	6.43	0.33
Permanently dropped out of the longitudinal sample during the 6-month wave (e.g. death of child before the age of 6 months, child separated from biological mother and moved to unknown location, permanent non-response, permanent loss of contact details)	138	32.86	1.67
Missed the 6-month wave, but did not drop out permanently (e.g. uncertain non-response, family problems, could not be interviewed in time due to change of primary carer, change of contact details or change of health visitor, technical organizational problem, lost interview)	255	60.71	3.08
Total	420	100	5.07

This gives a drop-out rate of 5.07 per cent relative to the prenatal wave.

As the original prenatal database in the figure above does not include the two drop-outs mentioned above, the number of subjects is 8,285.

Length of the interviews

The main questionnaire asked of the biological mother took an average of 38 minutes to complete (SD=19 minutes); in addition there was the consent form (if required) and the self-administered questionnaire to complete. As Figure 6 shows, quite frequently the health visitors entered full hours or half or quarter hours for the duration of the interview, in which case they probably did not enter the exact time it took to complete, but instead a 'rounded' value.



Statistics

		Lenght of the interviews
N	Valid	8340
	Missing	23
Mean		0:38
Median		0:35
Mode		0:30
SD		0:19
Minimum		0:03
Maximum		6:22

Source: Cohort '18 study 6-month wave, unweighted database, n=8340 (children).

Figure 6

Number of interviews (vertical axis) by length of interview (hours:minutes), main questionnaire with biological mother

It took significantly longer to complete the main questionnaire if the interviewer was asked to fill it in on paper; if the interviewer had to provide some form of assistance in completing the self-administered questionnaire (meaning that the respondent may interpret the questions with limitation); if the interview was interrupted (for a shorter or longer period); if the respondent was restricted in responding by any limitations (hearing, sight, speech, etc.); if the questionnaire was administered not in the office of the health visitor, but in the mother's home (or elsewhere); in the case of twin interviews; if the mother had at most 8 years of primary education; or if she had four or more children *(Table 3)*.

The questionnaire for primary caregivers (non-biological mothers) was completed by 29 respondents; it also took an average of 38 minutes (SD=17 minutes) to complete. The length of interviews (due to the small number of cases) did not differ significantly by background variables.

Table 3

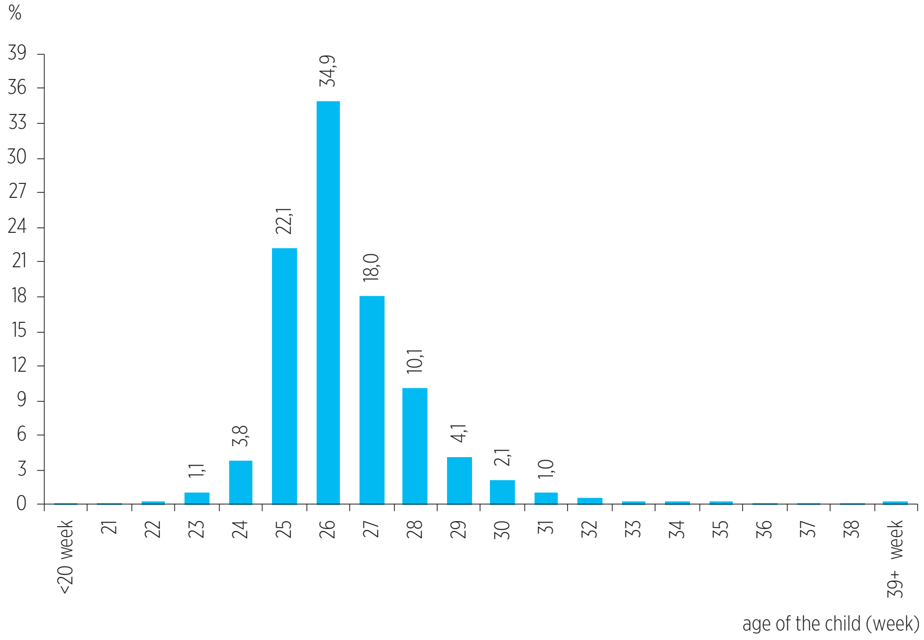
Length of interview by given key variables, main questionnaire with biological mother (mean, SD, N)

	Length of interview by given categories (mean, SD, N, significant differences between means)			
	Mean	SD	N	Welch's test, sign.
Total sample, n=8335	0:38	0:19	8335	
Interview method				****
Paper-based	0:41	0:19	6290	
Online	0:32	0:15	2046	
Assistance in completing the self-administered questionnaire				****
Completed independently	0:37	0:17	6971	
Asked for assistance with some questions	0:42	0:19	719	
Health visitor read the questions; respondent ticked answers	1:11	0:49	52	
Health visitor read the questions and ticked answers	0:43	0:28	519	
Respondent did not complete the self-administered questionnaire	0:43	0:21	74	
Interruption of the interview				****
No, was completed all at once	0:36	0:18	6060	
Yes, but only for a short period	0:43	0:20	2257	
Yes, for a longer period (several hours, days)	1:13	1:01	18	
Respondent's limitation				****
Limited (vision, hearing, other)	0:46	0:18	347	
Not limited by anything	0:38	0:19	7988	
Location of the interview				****
In the home of the expectant mother	0:40	0:19	5155	
In the health visitor's office	0:35	0:18	3061	
Other location	0:43	0:15	119	
Twins				****
One child was born	0:38	0:18	8103	
Twins were born	0:53	0:23	228	
Triplets were born	1:00	0:00	3	
Child's developmental delay				****
The status check has not yet been done	0:39	0:16	405	
No developmental delay was identified	0:38	0:19	7339	
Developmental delay was identified	0:39	0:18	456	
Health visitor did not know whether there's a developmental delay	0:44	0:12	134	
Educational attainment of mother				***
At most 8 primary school years	0:40	0:19	1656	
Vocational training	0:39	0:18	965	
Secondary	0:38	0:18	2817	
Tertiary	0:38	0:19	2890	
No response	0:43	0:16	8	
Age of mother at the time of the interview				ns
13–25	0:39	0:19	1501	
26–34	0:38	0:19	4662	
35–49	0:39	0:18	2173	
Actual partnership status				ns
Married	0:38	0:20	4785	
In a civil partnership	0:39	0:17	3218	
LAT or single	0:40	0:19	331	
Number of live-born children				****
1	0:37	0:18	3775	
2	0:39	0:20	2777	
3	0:39	0:19	1186	
4+	0:42	0:21	592	

Notes: p sign: ****<0.000; ***=<0.005.
Source: Cohort '18 study 6-month wave, weighted database.

Age of the child at the time of the 6-month interview

The second wave of the Cohort '18 study took place when the child was 6 months old. The 'target' range was 6 months minus 1 week/plus 3 weeks. The majority of interviews – covering 89 per cent of children – fell within this range: the average age of the children at the time of the interview was 26.5 weeks (with both median and mode values of 26 weeks); 5 per cent of children were interviewed at 24 weeks or earlier, while another 5 per cent were interviewed at 30 weeks or over. In the case of delay, the reasons most often cited by the health visitors were the respondent having moved house and illness (whether of the mother/child or the health visitor). According to health visitor comments, earlier completion was usually due to 'family going away for a longer period of time' or 'child going into hospital', or else because of family commitments.



Source: Cohort '18 study 6-month wave, unweighted database, n=8393 (children).

Figure 7
Distribution of children by age at the time of the 6-month survey

Data from the health visitor questionnaire

After the main questionnaire was completed, a so-called 'health visitor questionnaire' was filled in, recording the circumstances of the interview. Three quarters (75 per cent) of all interviews with biological mothers were conducted by the health visitors using a paper questionnaire, with subsequent web input; meanwhile one quarter (25 per cent) were recorded directly online during the interview. (Eight months earlier, in the prenatal interview, the split was two thirds and one third.) Most of the interviews were conducted in the mother's home (62 per cent), with a smaller proportion in the health visitor's office (36 per cent). This was also a change from the prenatal-wave interview, when more than half of the interviews were conducted in the health visitor's office. In 123 cases (1.5 per cent), the health visitors indicated other locations, such as the health visitor's home, a family support centre, or the home of the respondent's parents.

Both in the manual for health visitors and during the training sessions organized for them, we emphasized that interviews should be by appointment, since it is important for the interviews to proceed smoothly. Nevertheless, in two thirds of cases, the mothers had someone else present during the interview: mostly children under the age of 6 (59 per cent of cases – most often the child referred to in the study, but not necessarily just that child), but a relatively high proportion of partners were also present (10 per cent of cases). In more than half of the interviews (55 per cent), only one person was present during the interview *(Table 4)*.

Table 4

People present during the interview: who was present (percentage of mentions) and how many people were present (distribution of respondents), biological mother main questionnaire

	Who was present (percentage of mentions, multiple choice)
Child under 6 years old	59.0
6+ years old child	3.1
Spouse, cohabiting or other partner	9.6
Other relative	5.8
Other adult	2.9
No one was present	33.6

	How many people were present (distribution of women)	
	N	%
No one else was present	2768	33.1
One other person was present	4572	54.7
Two other persons were present	924	11.0
Three other persons were present	89	1.1
Four other persons were present	10	0.1
Total	8363	100

Source: Cohort '18 study, unweighted database, n=8363 children.

Almost three quarters (73 per cent) of maternal interviews went ahead without any interruption; 27 per cent had to be interrupted for a short period; and in 18 cases the interview had to be continued several hours or even days later.

The majority of mothers (96 per cent) had no difficulty in responding *(Table 5)*.

Table 5

Whether the respondent was limited in responding, percentage of mentions, more answers possible, biological mother main questionnaire

	The number of mentions of given limiting obstacles, multiple response options
Poor eyesight	8
Poor hearing	15
Did not understand (or had difficulty understanding) the question	161
Did not understand Hungarian well, language difficulties	16
Other	134

	Limiting factors mentioned, distribution	
	N	%
No limiting factors	8035	96.1
One limiting factor	324	3.9
Two limiting factors	2	0.0
Three limiting factors	2	0.0
Total	8363	100

Source: Cohort '18 study, unweighted database, n=8363 children.

Of the possible limiting factors, the most common issue reported by health visitors was that the mother did not understand (or had difficulty in understanding) some of the questions (161 cases). For those with language difficulties, either the partner or the health visitor herself helped the mother to understand the questions. The other most common limiting factor was care of the child (or children) and having to deal with them: attending to their crying, changing their nappy, feeding them, dealing with a tantrum or having to cook. As in the case of the prenatal-wave interview, three mothers reported mental health problems.

In most cases (85 per cent), the *self-administered questionnaire* was completed by the mothers themselves; only 8 per cent needed help with some questions. In 6 per cent of cases, the health visitor asked and completed the self-administered questionnaire herself (this was permitted if the mother asked the professional for help).

At the end of the health visitor questionnaire, the health visitor also had to indicate whether there was any sign in the 6-month-old child of developmental delay, based on the 6-month parental questionnaire and the health visitor status check assessment. In most cases (88 per cent) there was no delay; in 5 per cent there was a delay; and in a further 5 per cent this type of status check assessment had not yet been undertaken (furthermore, two health visitors either could not or did not want to answer this question).

Of the *primary caregiver (non-biological mother)* interviews, 69 per cent were recorded on paper and 31 per cent were directly input online. Most respondents were interviewed at home (55 per cent), while 28 per cent were interviewed at the health visitor's office; in five cases we do not know where the interview took place. In 41 per cent of cases, no one (other than the interviewer and the interviewee) was present during the interview, but in 59 per cent of cases someone else was present – mostly another adult(s) (six instances). Almost all interviews (90 per cent) could be conducted by the interviewer at one sitting, and in only three cases did the interview have to be interrupted for a short time. When asked what (if anything) had hindered the respondent during the interview, in one case the health visitor indicated that the respondent had not understood (or had had difficulty in understanding) some of the questions (and in five cases we do not know if there was any constraint). From the primary caregiver questionnaire and the health visitor's status check assessment, 20 children showed no indication of developmental delay; two children had some developmental delay, and in two cases, no assessment had yet been conducted on the children (in five cases there is no answer on this question).

2.4. COMMUNICATION ON THE DATA COLLECTION

During the 6-month study wave, we continued to use the communication interfaces developed at the start of the prenatal wave (Szabó et al., 2021). Our main goal was to maintain the participants' readiness to respond and to deepen their engagement regarding the research. These objectives were served by a colourful information booklet prepared for respondents, which was sent out together with the questionnaire packs and distributed to participants by the health visitors. In addition, information was made available via online channels (an official Facebook page and a website), as well as through stories in the media and other publications.

A colourful information booklet was included in the 6-month wave questionnaire pack and was handed to the mother by the professional when the details of the 6-month interview were discussed. This booklet again summarized the main points and stages of the study and the implications of taking part in the research. In addition, infographics were used to present some of the interesting findings from the prenatal survey wave.

For mothers who had not participated in the prenatal wave and only joined the survey when their child was 6 months old, a privacy and data-protection information leaflet was included as part of the questionnaire pack. This leaflet provided newly participating mothers with a detailed explanation of the purposes to which we put their personal data, the legal safeguards under which we use it, and how we guarantee the anonymity of participants during the data processing.

The research website (www.kohorsz18.hu), which was launched ahead of the prenatal wave, continued to provide participants with information about the study and supply all the contact details they needed to reach the research team directly if they had any questions. In addition to useful information on the data collection, the website also provided some brief analyses of the preliminary results of the 6-month wave (*Who does the housework?, Grandparent involvement II, Nurturing babies, Diary of the 6-month-olds, Partnerships II*). In addition to the short analyses, links to papers on the results of the research are also available on the site. The website likewise offers an opportunity to sign up to the newsletter that is sent out roughly quarterly to keep participants informed of the progress of the research.

In addition to the study's website, respondents can visit the official Cohort '18 Facebook page to find out more about the ongoing stages of the study, preliminary results, media coverage and other topics related to the research. Alongside all this, between August 2020 and January 2021, a feature ran on the Facebook page, every fortnight introducing one of the researchers working on the survey.

3. DATABASES

3.1. QUESTIONS SURROUNDING LINKAGE OF THE DATABASES

The research sub-databases for the 6-month study wave include data from the main questionnaire, the self-administered questionnaire and the health visitor questionnaire. The longitudinal database containing the data from the first two waves of the research will be created by linking these with the prenatal-wave databases.

The linking of each research database to each other and to the databases produced during the prenatal wave is done using the token, which is the identifier of the child (first-born child in the case of twins) and the mother/primary caregiver as the respondent.

The following variables help to navigate across the data in the longitudinal database:

t1_K1: Are the data from the main questionnaire for pregnant women available? 0: no (i.e.: joining at the 6-month stage), 1: yes

t1_self: Did someone fill in (at least part of) the self-administered questionnaire for pregnant women? 0: no, 1: yes

t1_K1idegen: Did the pregnant woman respond in a foreign language? 0: no, 1: yes

t2_K1: Is the data from the 6-month main questionnaire available from the birth mother? 0: no (i.e.: no birth mother version was completed), 1: yes

t2_K1proxy: Has a proxy questionnaire been completed in lieu of the prenatal-wave questionnaire? 0: no, 1: yes

t2_K2: Are the data from the main questionnaire at the 6-month stage available from a non-birth mother? 0: no (i.e.: birth-mother version was completed), 1: yes

t2_self: Has someone completed (at least partially) the 6-month self-administered questionnaire? 0: no, 1: yes

The sample sizes of the research sub-databases and their individual items are summarized in Table 6.

Table 6

Sample sizes of research databases

Name of the tool	Subset	Sample size: mother	Sample size: child
6-month main questionnaire	Proxy questionnaire to replace the prenatal-wave questionnaire	383	388
	Birth-mother version	8241	8363
	Primary caregiver version	29	29
Health visitor questionnaire	Birth-mother version	8241	8363
	Primary caregiver version	29	29
Self-administered questionnaire	Birth-mother version	7980 (at least partially completed questionnaire)	8092 (at least partially completed questionnaire)

3.2. DATABASE OF THE 6-MONTH MAIN QUESTIONNAIRE

As with the main questionnaire for pregnant women, the main questionnaire at 6 months could be completed by the health visitor in two ways: by recording the mother's verbal answers directly on the web interface, or by completing the questionnaire on paper and later inputting the answers on the online interface. As well as all the answers to the questions this database contains the most important data stored on the front of the main questionnaire: the identification of the child(ren) and the health visitor, the health visitor district identification, any case of twin pregnancy, the mother's year and month of birth, the date of the interview, the time of the interview, whether the respondent had been interviewed during pregnancy. In addition, the respondent is assigned a serial number in the database, which is an unmodified characteristic recorded in the online interface.

In the 6-month study wave, two versions of the main questionnaire are distinguished. If the child was being cared for by the birth mother at the age of 6 months, that birth mother answered the full main questionnaire (n=8241 mothers, n=8363 children). However, if the child was not being cared for by the birth mother, a shortened and modified version of the main questionnaire was answered by the child's primary caregiver (n=29 caregivers, n=29 children). These two types of main questionnaire are found in a common 'child-level' database (n=8392). However, it should be noted that in the case of other primary caregivers, there is a higher proportion of missing data, both cross-sectionally (due to the abbreviated questionnaire) and longitudinally (lower linkability to the pregnant women), than if it was the birth mother who responded.

If the mother, although part of the sample, had not previously been interviewed as a pregnant woman, the main questionnaire was preceded by a proxy questionnaire (n=383 mothers, n=388 children). Responses to this proxy questionnaire were recorded as 157 variables. The majority of questions were closed ended, but some questions required a number or date to be entered as an open-ended response. We also plugged the data content of the proxy questionnaire back into the corresponding variables of the prenatal database in the linked longitudinal database. Thus, for example, the mother tongue indicated in the proxy questionnaire is included not only in the variable corresponding to the proxy question, but also in the mother-tongue variable of the prenatal questionnaire. At the same time, data recorded during pregnancy but not included in the proxy questionnaire are, by definition, missing from the longitudinal database for those who joined later, at the 6-month stage.

The responses to the main questionnaire with the birth mother were recorded as 272 variables. The majority of these were also closed questions or in number or date format. The questionnaire also contained seven open-ended 'other' response options and six fully open-ended text-response questions (type of institution where the mother lives if not in a private household; reasons for child's hospitalization and medical care for injury or accident; first food tasted by child; reasons for mother and child spending nights apart). In addition, three new variables were created from the questions on infant temperament (see subsection 7.2).

In the case of twins, the caregiver also answered the questions of the main questionnaire for the second child (except for the question on the sharing of childcare responsibilities (F37)). The data on twins belonging to the same variable are included in the longitudinal database as the same variable, but on separate rows. (By implication, the token identification variable is the same for twins: only the child identifiers – child_ID – are different.)

Based on the need assessed in advance, for the pregnancy wave we had prepared several foreign-language versions of the questionnaire (English, German, Chinese and Vietnamese) (to be self-administered). However, the uptake of these self-administered questionnaires was very low, and thus, no foreign-language version of the questionnaire was prepared at all for the 6-month survey. Six mothers who had previously filled in a foreign-language version answered the main questionnaire in Hungarian. One of them answered the Hungarian questionnaire without assistance and without difficulty, while the other five were helped to understand the questions either by the health visitor or by a family member.

Responses to the main questionnaire with a primary caregiver other than the birth mother were recorded as 145 variables, including 13 text responses. Wherever possible, the data content of this version of the questionnaire was also plugged into the variables created for the birth-mother questionnaires. For example, the response given by the primary caregiver on the general health of the child was included not only in the variable for primary caregivers (labelled K2), but also in the variable for birth mothers (labelled q2).

3.3. DATABASE OF THE HEALTH VISITOR QUESTIONNAIRE

The questionnaire that was completed by the health visitors via the online interface after the interview asked about the circumstances of the interview and whether the child in the study had any developmental delay according to the health visitor's status check. The data from the questionnaire were recorded by the health visitors on the online platform using the token, so that the 26 variables recorded automatically became part of the main questionnaire database. The questionnaire contains mostly closed-ended questions, plus five open-ended text explanations and one fully open-ended 'other' comment field.

3.4. DATABASE OF THE SELF-ADMINISTERED QUESTIONNAIRE

Questions on certain more sensitive topics – such as the mother's emotional and psychological state, the birth, nationality and religion – were asked using self-administered paper-based questionnaires. These were given to the mother by the professional in the form of a booklet after the main questionnaire had been completed. The self-administered questionnaire was sealed in an envelope immediately after completion and sent by the health visitor to our institute, where it was recorded. Because the self-administered questionnaire was given only to birth mothers, the maximum possible number of respondents was 8,241; of these mothers, 7,980 completed at least part of the questionnaire.

The self-administered database also contains a token (which can be used to link to the main questionnaire data and to link the prenatal-wave database to it), a serial number, the child(ren)'s ID and the mother's month and year of birth. The questions in the self-administered questionnaire were recorded as a total of 110 variables. Most of these were also in closed-ended response format or in number format, but they also included 11 open-ended answers, one open-ended textual justification and one completely open-ended answer (municipality where the mother gave birth). Five new variables were created from the questions in the self-administered questionnaire (see subsection 7.2) and are also part of the database.

3.5. DATABASE OF CONSENT FORMS

If a new respondent joined the study at the 6-month stage, a consent form had to be completed – either by the birth mother (who for some reason had been unable to respond at the time of her pregnancy) or by another person who was now the child's primary carer. The consent form details were recorded at our institute: the respondent's birth name, her mother's name, the place and date of the respondent's birth, and her actual and official address. The respondent also had the option of providing a telephone number, email address and social security number. These data are stored separately from the questionnaire responses, in a secure, password-protected file. In accordance with data-protection laws, it is not accessible to third parties or to unauthorized staff at HDRI. When necessary, for verification purposes the mother's date of birth was compared to the same data from the 6-month main questionnaire.

3.6. ADDRESS DATA

At the very beginning of the study, in the prenatal wave, the health visitors were given a folder in which to collect the main data from the interviews successfully completed during the prenatal and infancy waves of the research, the number of unsuccessful interviews and some of their main characteristics. (In the event that the health visitor in a given district was unable to continue the interviewing, the folder was taken over by the health visitor who replaced her.) These folders were collected at the end of the fieldwork and their data content was recorded by HDRI staff on a database stored separately from the survey data on a password-protected computer.

As the research progressed, the health visitors also gathered in this folder the respondents' contact and recruitment data from the index cards. This information was handed over to our institute at the end of the 6-month fieldwork. The content of the data on the index cards was then compared by HDRI staff with the data in the address database, which is constantly updated; any relevant changes were recorded.

4. DATA EDITING AND CLEANING

4.1. EDITING UNIQUE IDENTIFIERS (TOKENS)

As in the prenatal wave, the online data-recording process for the 6-month wave was designed in such a way that only one questionnaire could be recorded and finalized with one case identifier (token). The system made all tokens that had been created, but not yet used, available to any health visitor, so that the continuity of respondents' unique identifiers could be maintained, even if the child moved to a different district or if the health visitor had changed for some reason since the previous data collection (e.g. job transfer, retirement, illness). In addition, the online system allowed the child identifier that was entered manually to be different from the token (the token was used to log in to the recording interface). In the case of twins / triplets, the second / third child identifiers were linked to a token, and only the first child identifier has to match the token.

This repetition of numbers served to ensure continuity and the linkage of cases: if an incorrect token was accidentally used by the health visitor at the login, the difference in the second number entered manually was likely to draw attention to a possible identification problem.

Errors affecting continuity and longitudinal database linkage could occur for a variety of reasons, not all of them associated with a mismatch between these two identifiers. A significant proportion was found when we compared the mother's date of birth by survey waves, or the estimated and actual date of birth of the child. Sometimes it was simply because the 6-month interview token was not linked to either the prenatal-wave databases or the consent form or address list data received at the time of the 6-month wave. Identification and continuity errors that occurred during the logging of the 6-month data were flagged up in the following ways.

Some interviews assumed a pregnancy-phase interview (recorded without a prenatal proxy module) that was not linked to any prenatal-wave questionnaire by a token.

In a number of cases, the recording of a 6-month interview with a mother who had already been interviewed in the prenatal wave encountered the problem that the token in question had already been used, i.e. another interview had previously been recorded with it by mistake. Some of these issues became apparent when it was necessary to use the wrongly ascribed token. However, on quite a few occasions, the health visitors who had made the incorrect recording immediately noticed their mistake, and they themselves indicated which token we should actually use. In some cases, they even attached the correct token, creating a duplicate.

It also sometimes happened that when we found we were missing an interview that was due, the health visitor reported having already prepared and logged it, though we could not find it by the token.

Comparison of the mother's date of birth between the two study waves revealed a number of errors, where the 6-month interview identifiers were not linked to the correct pregnancy history because two or more cases had been logged with reversed numbers.

For the interviews recorded with the prenatal-proxy module (i.e. without a prenatal-wave history), the consent forms and address data were obtained during the 6-month wave, as appropriate. But certain difficulties were also encountered in this area: some interviews were not accompanied by the consent form marked with the appropriate token by the required date; meanwhile, in other cases there were delays in recording the interview online and there was no way of associating the consent form received beforehand. In order to obtain the missing records and consent forms, we obviously needed to know what was lacking and whether it was a real data gap or just a mismatch.

Thus, the accurate and reliable identification of cases was very important for several reasons. Issues in this area may have arisen from simple oversights or inadvertent mix-ups – especially when inputting several paper interviews at the same time – but also from occasional misinterpretation of the counting methodology. Sometimes, the health visitor took the token not as the ID number of the respondent, but as the ID number of the interview, and therefore initially logged the 6-month questionnaire using the next ID number. Moreover, some of the health visitors who joined the research at a later date reassigned their own four-digit ID to cases they had taken over from their predecessor (i.e. they continued with different numbers). A single accident or mix-up could have a domino effect involving up to 4–5 cases. However, we were given every opportunity to clarify complex situations, because if we did not reach a satisfactory solution by reviewing and collating the data available to us, we could go back to the field and ask the health visitors, who knew their interviewees and could therefore sort out the problem on the basis of the information transmitted. The main support in resolving identification problems was therefore provided by comparison of the data from the two study waves: first, the correspondence between the mother's date of birth and secondly, the coherence between the expected and the actual birth dates of the child. The assistance of the health visitor was sought when the data available did not adequately allow for the unequivocal identification of cases. Subsequently, however, we were able to clarify doubtful circumstances with maximum efficiency, so that no 6-month interviews were lost due to problems of identification.

We corrected the token or the child identifier (or both) in 192 cases and deleted eight cases from the original database of 8,250 birth-mother interviews: seven because they were duplicates (due to logging errors) and one because we did not receive the consent form and were unable to obtain it subsequently.

Of the 385 cases recorded with the proxy module, 383 new entrants at the 6-month stage were ultimately included in the study. One case was omitted due to the aforementioned lack of a consent form, while the other had a clearly identifiable prenatal-wave record (i.e. the prenatal proxy questions were asked by mistake at the time of the 6-month interview). However, there are five interviews in the 6-month database that lack a proxy module, despite having no history in the prenatal database. In two of these cases, a pregnancy-wave interview was conducted, but was not included in the analysis database because it was very incomplete; in the other three cases, the proxy module was simply missing.

A total of 8,085 cases were recorded in the self-administered database: this is the number of self-administered questionnaires returned to HDRI that had been completed (to varying degrees). As a first step, we checked the identifiers (the tokens) in the self-completion database against the mother's date of birth and corrected any incorrect identifiers, if the data allowed. After the identifiers were corrected, the records with incorrect or duplicate identifiers that could not be corrected were deleted (89 cases). We then linked the self-administered database to the main questionnaire database and deleted five additional cases that could not be matched to main questionnaire and eight cases that belonged to a non-biological mother (since the self-administered questionnaire was not to be completed by other caregiver than biological mother).

4.2. CLEANING THE VARIABLES

The biological mother primary caregiver questionnaire: main interview

The checking and cleaning of the biological mother primary caregiver database of the Cohort '18 study's 6-month survey was carried out in a way similar to the prenatal wave.

Since most of the questions were close-ended, the database frame could be prepared in advance using a data-entry program. The data were entered into the database by the health visitors using this pre-written recording program. Thus, a filtering program was already in place during the recording process. That is, the health visitors could not record any value that was not included in the recording program, which was compiled by the researchers and the survey programmers on the basis of the questionnaire response categories and possible response intervals. The pre-designed recording program also left blank those questions that were skipped.

The databases were checked both at the level of variables and at an individual record level. As a first step, the frequency distribution of each variable was checked. If there were values that could not have been present, they were listed at the individual record level, checked and corrected (if necessary). Such errors could occur if a respondent realized during the interview that she had given an incorrect answer to a previous question, so that the interviewer had to go back and overwrite the previous incorrect answer. In such cases, the incorrect answers were left on the 'incorrect path' and had to be deleted afterwards.

Logical relationships and inconsistencies were checked and cleaned at the *cross-sectional level* based on the responses to the 6-month survey. That is, even if there was a discrepancy between, for example, a response during pregnancy and a response at 6 months, the data at 6 months were accepted as valid and were not overridden by the data at pregnancy. For example, if a pregnant woman had been living with her husband at the time of the prenatal wave, but now during the 6-month survey she said that her relationship status had not changed since the previous wave and that she was living in a cohabiting partnership, we kept as 'true' the 6-month response. Or if at the time of the prenatal wave the mother-to-be said that her partner was the biological father of the unborn child, but at the 6-month interview she said that that same partner was not the biological father of the child, we also kept as 'true' the second-wave response. Such data inconsistencies occurred in a small number of cases: inconsistencies related to parental partnership history occurred in about 3 per cent of cases, while inconsistencies related to changes in the status of the biological father occurred in just under 21 cases. These statistical/logical inconsistencies in the database may in some cases even be true to life. There may be periods of time when the partnership situation changes so often that on the day of the interview, even though the respondent considered a statement to be true for her partnership situation, it was not the case on the previous day, or she knows it will not be the case the following day, the following week or the following month. The acknowledgement of the biological father status could also change over time, as the mother may be more likely to call her partner the biological father of the child, even if she admitted and acknowledged during pregnancy that her partner was not the biological father of the child.

In the case of the table-format multi-response questions (A02, F09/G09, F21/G21), there was always a response category 'none of the above'. If, for a given question, the respondent answered 'none of the above', but also answered 'yes' to a specific question in the table, the specific answer was taken as true. For example, for the question (F09/G09) on what medical conditions the child had, a respondent may have mentioned hip problems, may have mentioned the hearing loss, and may have mentioned also "none of these". In such cases, the answer to the question 'none of these' was deleted.

For questions where numbers (e.g. weight, length/height, size of dwelling, estimated value of dwelling) had to be entered, in many cases the recorders also wrote text next to the numbers. In such cases, the text was manually deleted to convert the variables into numerical or date format.

The biological mother primary caregiver questionnaire: proxy questionnaire on pregnancy

Mothers who, for whatever reason, could not be interviewed during the prenatal wave were subsequently given an opportunity to answer a so-called prenatal proxy questionnaire at the beginning of the 6-month survey. The questions also form part of the 6-month dataset (but are not included in the processing of the questions on the 6-month data). However, the answers to the retrospective proxy questions were entered into the pregnancy database, so when we do a longitudinal analysis using data from the prenatal and infancy (6-month) waves, we have answers for both time points in these cases as well.

The proxy pregnancy-related variables were checked and cleaned in a similar way as the main questionnaire variables. We checked for skips and data gaps based on the variable distributions. Since the health visitors completed a pre-programmed questionnaire, there were no invalid values and the health visitors could not circumvent the skips. As mentioned earlier, there was one case where there was both an original prenatal-wave questionnaire and a 6-month prenatal proxy questionnaire; here, we accepted the data from the prenatal wave and ignored the information from the proxy questionnaire. We found inconsistencies in only two cases: two mothers did not provide data on the total number of live children they had (for technical reasons, one of them was only able to record 10 of her 14 children). The data on the value of the dwelling sometimes had to be edited, because not all the women gave the value in millions of forints and sometimes they gave the answer in text form (13 cases).

The non-biological mother primary caregiver questionnaire

We interviewed a total of 29 primary caregivers who were not the biological mother of the 6-month-old child: most were professional foster parents (12 cases) or adoptive parents (11 cases). The remaining six respondents were grandmothers (or in one case an aunt).

Overall, we received a very coherent and error-free database from the health visitors. The questionnaire contained few skips, but not everyone had to answer all the questions (minimum 122 and maximum 127). As the questionnaire was recorded by the health visitors on a pre-programmed interface, there were no non-responses (or over-responses) and no response codes that were not included in the questionnaire.

The biological mother primary caregiver questionnaire: self-administered questionnaire

The cleaning of the self-administered database was done by reviewing the distribution of the variables and filtering out the erroneous values. As the self-administered questionnaire was completed on paper by the participants and their answers subsequently recorded by a person assigned by HDRI, a significant proportion of the data cleaning was due to manual recording errors. The cleaning thus consisted of correcting or deleting responses that were not possible on the given scale. In addition, for certain free text and ‘other’ answers to some question sets (negative life events, place of delivery, other person present at delivery, other intervention during birth), some logical editing was necessary. In fact, some of the text responses in the category ‘other’ could be (or already were) classified in one of the predefined categories, so they were marked there and ‘other’ was deleted. For some ‘other’ answers (name of the settlement of residence, religion), typos and spelling errors were corrected. Finally, for the question on interventions during birth, we deleted ‘no intervention’ if the mother indicated an intervention (six cases).

5. DATA QUALITY CHECK

5.1. MONITORING THE INTERVIEWS

Monitoring of the 6-month data collection was continuous. As a first step, a spreadsheet was compiled, listing the most important information used in monitoring for a given case identified by a token. The following background data were used for verification: token, the health visitor district identifier, week of pregnancy at the time of the prenatal-wave interview; type of pregnancy; foetal identifiers in the case of twins; the pregnant woman's date of birth; due date of delivery; date on which the questionnaire was complete; and the exact date on which the child will turn 6 months. In addition, we included information that showed whether the mother had moved house. All this information was used to determine the time window when the 6-month interview should take place. If the questionnaire data were not received by the expected date, we first contacted the health visitor by email; if she did not respond, we tried telephone. After a successful consultation, we followed up to see if the health visitor had recorded the missing questionnaire.

The completion of the 6-month survey of mothers could be delayed for a number of reasons:

The health visitor may have left her district (illness, change of profession, etc.) without HDRI having been informed; thus the institute could not contact the new health visitor to enrol her in the research. As soon as we learned of such cases, we immediately invited the new health visitor to participate in the research, signed a contract with her and explained the entire study and the interviewing process (i.e. trained her).

The respondent may not have been available at the time of the interview due to illness (of the mother or child) or some other serious family problem or other reason.

The family may have moved without the health visitor indicating the new address; in such cases, we requested the new address and told the district health visitor about the new address where a survey needed to be carried out.

The mother may have moved abroad or no longer wished to participate in the research, but the health visitor had failed to indicate this in the specified interface; after consultation, we closed the address with the appropriate status code.

In the case of queries that were completed on time, delays in the valid recording of the data series could sometimes occur, for a number of reasons:

The health visitor may have recorded the questionnaire, but had mistyped the ID (token), thus (initially) recording the data under an incorrect token.

The health visitor may have interviewed the mother, but not yet recorded the questionnaire.

In the former, a correction procedure was used to fix the discrepancy; in the latter, group emails were sent to the health visitors to remind them of the importance of recording, and there was continuous monitoring to address the problem. This was also for their own benefit, as the payments due for carrying out a survey were related not to the number of interviews conducted, but to the number of interviews recorded under accurate and cleared tokens.

5.2. CHECK FOR ITEM NON-RESPONSE AT VARIABLE LEVEL

The biological mother primary caregiver questionnaire: main interview

A total of 17 mothers were mentioned by health visitors as not understanding Hungarian well, but all of them had help from someone – usually the woman's husband, partner or the health visitor herself. So, the sometimes higher proportion of 'don't know' or 'don't want to answer' responses among respondents was not necessarily a consequence of their not understanding the language.

A relatively high proportion of 'don't know' and 'don't want to answer' responses was found when the mother had to answer questions about her *partnership situation*. When it was a question about an unchanged partner (i.e. the same as at the time of the pregnancy), a higher proportion of respondents indicated that they were unsure whether they planned to get married in the near future (12 per cent of cohabiting partners said they did not know). If they had married since the seventh month of pregnancy, 5 per cent and 8 per cent, respectively, did not know (or did not want to say) in which year and month they had done so, even though it must have been in either 2018 or 2019. Some 15 per cent of those cohabiting women who had not married yet did not know (or did not want to say) if they planned to. Relatively few mothers had had a new partner in the 8 months since the seventh month of pregnancy, but a relatively high proportion of them were unable or unwilling to give the year of birth of either their new cohabiting partner (three mothers out of 31) or their new living apart together (LAT) partner (six out of 28). A total of 278 children had a father living separately. The mother did not know the family status of the father of 21 percent of the children (and a further 5 per cent did not want to answer).

Among *homeowners moved to a new home*, 11 per cent of respondents could not estimate the value of the home, and 11 per cent could not estimate how many square metres it contained; and 15 per cent of those moving could not decide whether or not they had a garden.

The proportion of those who did not know or did not want to answer was also relatively high for questions regarding *employment*. The mother of 18% of the children did not know if she would be able to do it part-time if she would go back to work.

In the case of the child questionnaire, there are hardly any questions that respondents could not answer. The questions that the highest proportions of mothers could not answer had to do with the timing of breastfeeding (9 per cent) and whether they knew if they themselves had been breastfed by their mothers (5 per cent). Somewhat surprisingly, 5 per cent of the mothers of twins could not say whether they had identical or fraternal twins.

The biological mother primary caregiver questionnaire: proxy questionnaire on pregnancy

The highest proportion of 'don't know' responses came when the mother was asked about her weight at birth (23 per cent of children had a mother who answered either 'don't know' or 'don't want to answer'), how many square metres they had to live in (9 per cent and 2 per cent, respectively) and what they thought the market value of the property was (i.e. how much it could be sold for) (8 per cent and 5 per cent). There was also a relatively high proportion of mothers who could not say in what month they had moved in with their partner (8 per cent), the year of birth of their LAT partner (8 per cent), his level of education (8 per cent) and the year in which their relationship had started (8 per cent).

The non-biological mother primary caregiver questionnaire

Overall, there were only a few questions that the 29 respondents to this questionnaire were unable or unwilling to answer. Most respondents (12 children' caregivers) could not or did not want to answer the open question 'Why exactly is the birth mother not the child's primary caregiver?' Five primary carers of children could not or did not want to answer 'How many biological siblings does the child have with whom he/she does not cohabit?'; and five carers could not answer 'In what week of pregnancy was the now 6-month-old child born?' There are four children for whom we have no information on 'How many nights did the child stay in hospital immediately after birth?' A further 10 questions had one or two missing or non-responses.

The biological mother primary caregiver questionnaire: self-administered questionnaire

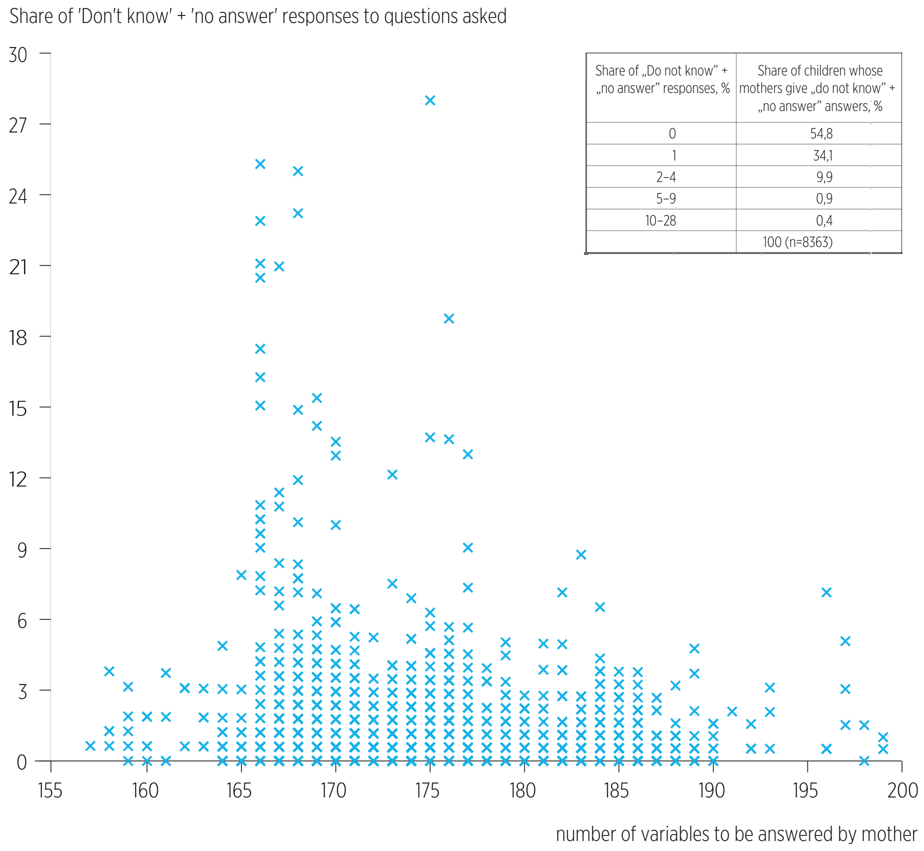
The self-administered questionnaire did not use 'don't know' and 'don't want to answer' as response options; and although respondents left some questions blank, these could not clearly be assessed as 'don't know' or 'don't want to answer' responses. However, if we consider the 7,929 mothers who answered at least one of the questions in the self-administered questionnaire as a baseline, the highest proportion of non-response was on the question about the amount of money spent on childbirth, to which 21.5 per cent did not respond. Similarly, a large number (20.1 per cent) did not answer the question on the number of nights spent in hospital before giving birth, presumably largely due to the missing response of 0 (the non-response rate for the number of days spent in hospital after the birth was only 2.2 per cent).

5.3. CHECK FOR NON-RESPONSE AT CASE LEVEL (UNIT NON-RESPONSE)

The biological mother primary caregiver questionnaire: main interview

The 121 questions of 6-month questionnaire were recorded as 272 variables. That is the maximum number of 'variables' that mothers would be asked if all questions were relevant. Of the 272 variables, 13 were from open-ended questions (mothers were allowed to give open text responses to a total of 20 questions, though these were only recorded as 13 variables, because the responses for twins are recorded as one variable in the child-level database). The remaining 259 variables were recorded as numerical variables. Taking these numerical variables into account, we looked at the percentage of 'variables' that mothers were unable or unwilling to answer. On the horizontal axis of Figure 8, we indicate the number of variables that the mothers had to answer (since not everyone had to answer all the questions: depending on their situation, mothers could answer a minimum of 157 and a maximum of 199 variables). On the vertical axis, we have indicated the percentage of questions that mothers could not or did not want to answer.

It is clear that the majority of the mothers answered all the questions, the share of 'don't know' and 'don't want to answer' is 0 for 54.8 per cent of the children. Furthermore, there was not a single respondent who answered 'don't know' or 'don't want to answer' to more than half of the questions (the maximum was 28 per cent). Mothers of one third of the children (34.1%) could not or did not want to answer 1 per cent of the questions; the mothers of 9.9% of the children could not or did not want to answer 2–4 per cent of the questions. In total, the mothers of 31 children (0.4 per cent) gave a relatively high proportion of ‘don’t know’ or ‘don’t want to answer’ responses: they were unable or unwilling to answer 10–28 per cent of the questions.



Share of „Do not know" + „no answer" responses, %	Share of children whose mothers give „do not know" + „no answer" answers, %
0	54,8
1	34,1
2–4	9,9
5–9	0,9
10–28	0,4
	100 (n=8363)

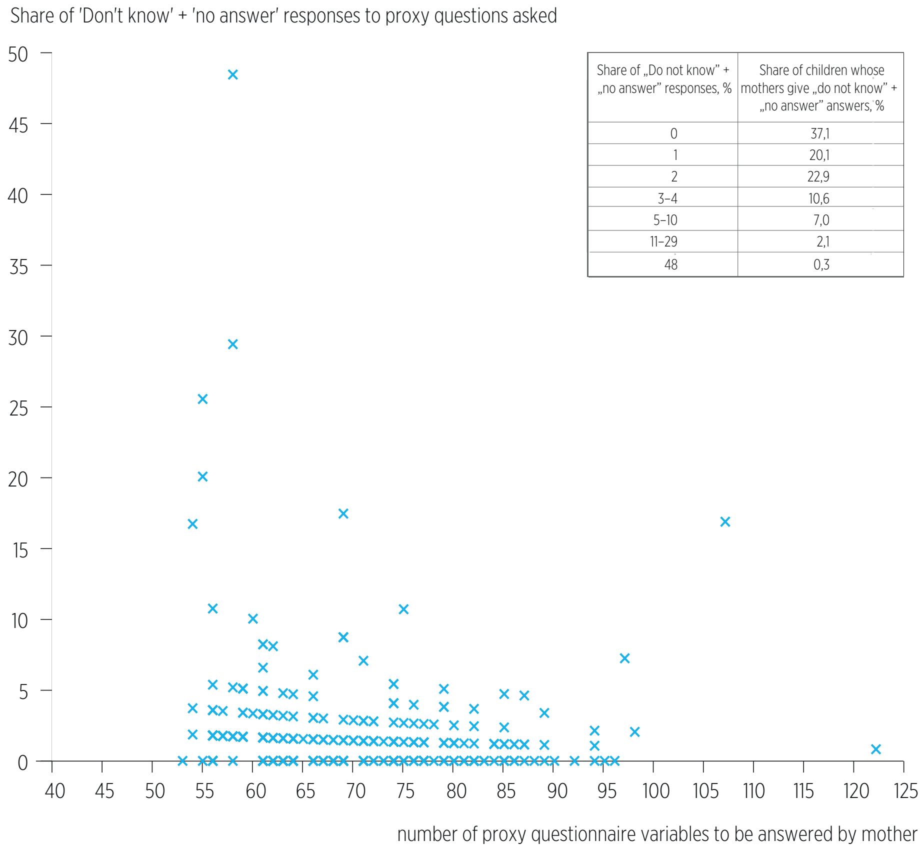
Source: Cohort ’18 6-month study wave, unweighted database, child level (n=8363). Crosses indicate one or more respondents.

Figure 8

Number of variable asked of the mother (x-axis) and the combined percentage of ‘don’t know/don’t want to answer’ responses (y-axis), main maternal questionnaire

The biological mother primary caregiver questionnaire: proxy questionnaire on pregnancy

The retrospective proxy questionnaire on pregnancy contained 48 questions, and the answers were recorded as a total of 157 variables. Two variables contained open-ended responses and the remaining 155 contained numerical responses. Naturally, in this case, not all proxy questions had to be asked of all mothers: depending on their situation, mothers had to answer a minimum of 53 and a maximum of 122 ‘variables’. Of those women who responded to the prenatal proxy questionnaire, 37.1 per cent answered all the variables asked (or more specifically, the mothers of 37.1 per cent of the children). There was one child whose mother did not answer almost half of the variables: she refused or was unable to answer 28 of the 58 variables asked (representing 48 per cent of the variables asked, see Figure 9). These refusals were due to the mother’s unwillingness to share information about her husband with the health visitor. For 10 children, the mother could not or did not want to answer 11–29 per cent of the variables (representing 2.1 per cent of children); for 27 children (7.0 per cent), the mother could not or did not want to answer 5–10 per cent of the variables; and for 41 children (10.6 per cent), the mother could not or did not want to answer 3–4 per cent of the variables. For nearly a quarter of the children, the mother was unable to answer 2 per cent of variables, while for a fifth of the children the mother was unable to answer 1 per cent of the variables *(Figure 9)*.



Share of „Do not know" + „no answer" responses, %	Share of children whose mothers give „do not know" + „no answer" answers, %
0	37,1
1	20,1
2	22,9
3–4	10,6
5–10	7,0
11–29	2,1
48	0,3

Source: Cohort '18 6-month study wave, unweighted database, child level (n=388). Crosses indicate one or more respondents.

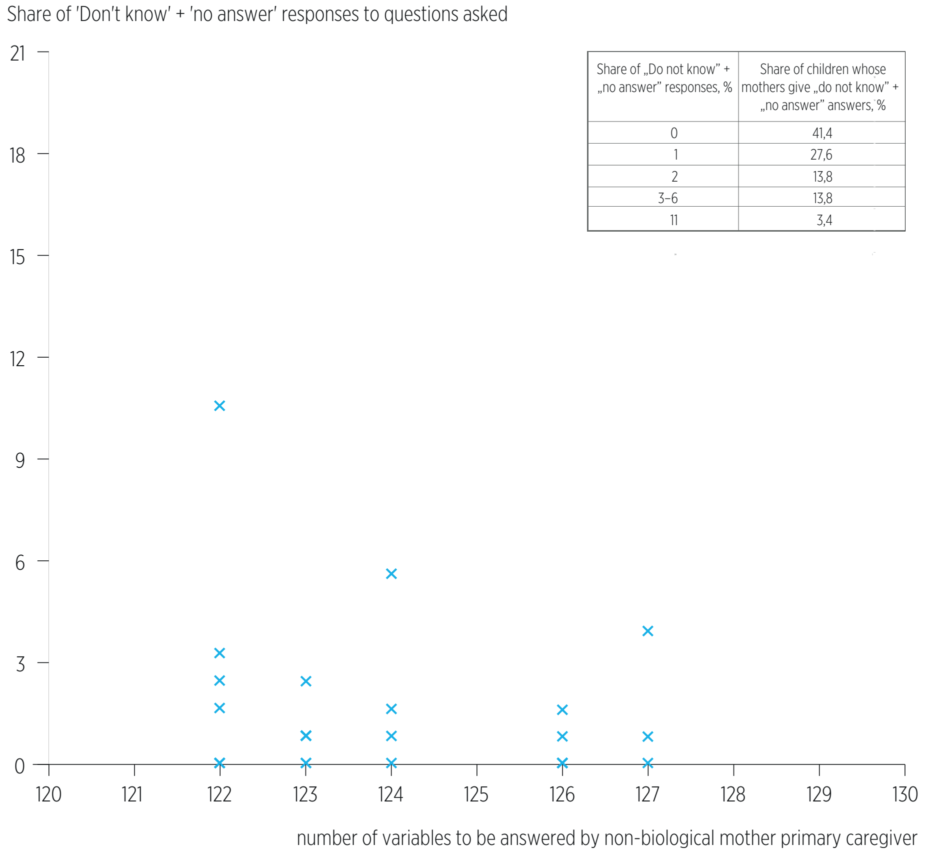
Figure 9

Number of questions asked of the mother (x-axis) and the combined percentage of 'don't know/don't want to answer' responses (y-axis), 6-month proxy survey

The non-biological mother primary caregiver questionnaire

Non-biological mother primary caregivers were asked a total of 64 questions about the child and about themselves and their family. The answers to the questions were recorded as a total of 145 variables: 132 numerical and 13 open-ended (most of which included an explanation of the 'other' response categories). Of the 13 open-ended questions, two had to be asked of everybody: (1) the exact reason why the child was not living with the birth mother; and (2) the occupation of the primary caregiver. It is true that the latter question was incorrectly answered: all but one caregiver failed to indicate an occupation, saying merely that they were currently professional foster parents. These variables are treated as numerical variables after automatic recoding. Respondents had to answer a minimum of 122 and a maximum of 127 variables *(Figure 10)*.

For the non-biological mother primary caregivers, 12 of the 29 respondents answered all the variables that were asked; eight could not or did not want to answer less than 1 per cent of the variables; five did not answer 2 per cent; three did not answer 3–6 per cent; and one respondent failed to answer 11 per cent of the variables. In this last case, the respondent could not or did not want to answer 13 out of the 122 variables asked (mainly concerning such sensitive questions as why the biological mother was not raising the child, questions related to the financial situation of the household and questions concerning her partner).



Share of „Do not know" + „no answer" responses, %	Share of children whose mothers give „do not know" + „no answer" answers, %
0	41,4
1	27,6
2	13,8
3–6	13,8
11	3,4

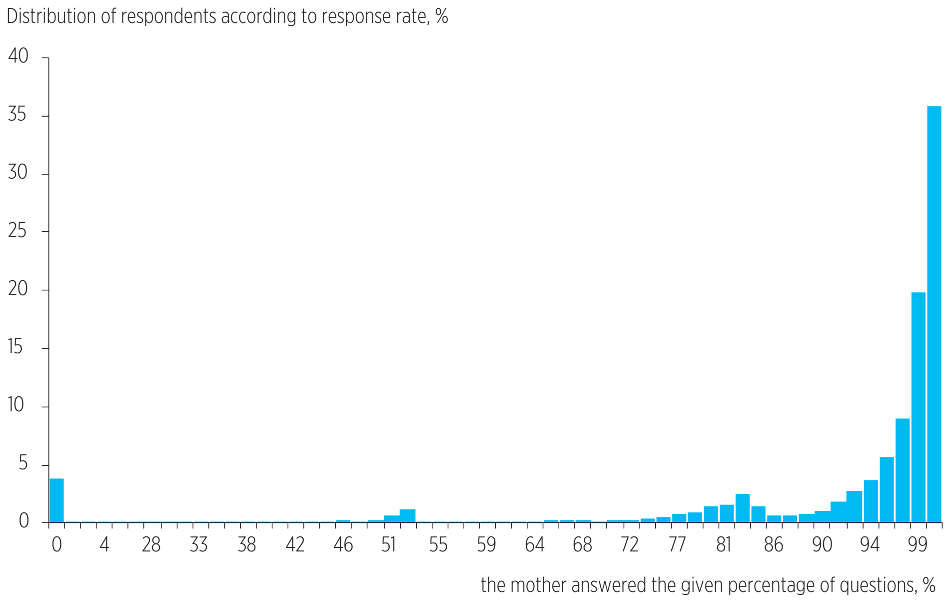
Source: Cohort '18 6-month study wave, unweighted database, child level (n=29). Crosses indicate one or more respondents.

Figure 10

Number of questions asked of the caregiver (x-axis) and the combined percentage of 'don't know/don't want to answer' responses (y-axis), non-biological mother caregiver

The biological mother primary caregiver questionnaire: self-administered questionnaire

All the data from the self-administered questionnaire were missing for 312 mothers – 3.8 per cent of respondents to the main (i.e. biological mother primary caregiver) questionnaire. They either did not fill it in at all or else their questionnaire could not be linked to their additional data due to lack of correct identification. No 'don't know' or 'don't want to answer' response options were used in the self-administered questionnaire, but respondents occasionally did leave some questions blank. There was no actual skipping here, but there were some questions to which we did not expect answers from all mothers (to 'other' response categories; to questions regarding the reasons for a caesarean section 4.9., and to partnership questions 7.1-7.2) or did not expect answers for all items of the questions (4.3, 4.8, 4.9). The theoretical minimum 'required response rate' was defined separately for single mothers (57 variables to answer) and mothers in a partnership (69 variables to answer), based on the response to the main questionnaire B04 on the actual partnership situation. Among mothers with partnership, 37.1 per cent answered every self-administered question. Less than 80 per cent of the variables were answered by 12.2 per cent; and less than 50 per cent by 5 per cent (including 3.8 per cent who answered no questions). The question on childbirth-related expenditure was answered by the lowest proportion of mothers: 75.5 per cent provided data on this.



Source: Cohort '18 6-month study wave, unweighted database, mother level (n=8241).

Figure 11

Distribution of respondents according to the proportion of respondents who answered the questions in the 6-month self-administered questionnaire

6. WEIGHTING

The purpose of weighting as a statistical procedure is to eliminate minor biases in the sample due to sampling, variations in response rates, selective attrition, other non-sampling errors, etc. This is done by matching the data from a given database as closely as possible to some 'real' (e.g. census population) or hypothetical population base. In the present case, the weights for the 6-month datasets have been generated as follows.

6.1. CREATION OF THE INITIAL LONGITUDINAL BASELINE WEIGHT (S_MAGZAT_V)

As a first step, a 'quasi-cross-sectional weight' was produced for the total number of the population participating in the research – the so-called 'total reference population of the cohort study' (for a definition of this, see Szabó et al., 2021). The function of this weight is to exactly match the aggregate data of all survey respondents to the population in terms of given distributions: e.g. the proportion of graduates among the birth mothers in our study should be equal to the proportion among all women who gave birth during this period.

The total population covered was 8,844 individuals: all the foetuses (and later children) and (associated primary) caregivers on whom data were collected during the study. The total population of concern is therefore the total number of children who could potentially participate in any of the data-collection waves of the research. Data on children who are not part of this population were not collected and will not be collected in the future.

The total cohort study population concerned is composed of several subgroups of different sizes. It is a mixed population, in the sense that not all of these 8,844 persons participated in both data-collection waves.

Of course, the largest group of the total population concerned is made up of those foetuses and infants who were enrolled in both the prenatal wave and the 6-month wave (both times involving a full Hungarian-language questionnaire): 7,968 cases were in this group – i.e. this group makes up the vast majority of the study population (90.1 per cent).

The second major subgroup consists of children who, for one reason or another (e.g. very premature birth, illness of the health visitor, etc.), were not enrolled during pregnancy, but whose mothers completed a replacement, so-called proxy, questionnaire when the infants were 6 months old (399 cases, 4.5 per cent). (Of these children, 388 were being cared for by their mother and 11 by someone other than their biological mother.)

The other groups that make up the remainder of the total affected population are significantly smaller: for example, those who participated in the prenatal wave but were 'lost' either permanently or temporarily because their family moved house or emigrated during either the prenatal or the infancy (6-month) waves; or if the caregiver refused to respond at the 6-month stage; if the child died; if an anonymous adoption took place; if a questionnaire was completed in a foreign language during the pregnancy wave, but this was no longer possible at 6 months, etc.

Minor inaccuracies due to sampling and differences in the response rates were reduced by first adjusting some data from this 8,844-item population to the 2018 HCSO Vital Statistics Register (VSR) of live births in Hungary, covering women with Hungarian residence or domicile. The weighting took into account the (estimated) partnership status of the mother at the time of birth, her educational attainment, her age group and the developmental level of the region of residence. This resulted in a special technical baseline weight (s_magzat_v), which formed the basis for the two weights to be used for the actual analysis. The technical details of the creation of this weight are identical to the method used and presented for the production of the weights for the prenatal child analysis database (Szabó et al., 2021). The reference population from the VSR was also the same, so that – for the individuals in both populations – the weights themselves were very similar. Consequently, the detailed technical description of the weighting procedure is not repeated here, as it can be found in the previous methodological volume.

There was only one problem that needs to be briefly addressed here: in some cases (31 cases, 0.4 per cent), at least one element of the mother's data for weighting the children was not known or was not relevant. For example, if the child was adopted after birth; or if the prenatal questionnaire was incomplete and then the mother refused point-blank to answer at the 6-month stage. In such cases, these children were assigned initial weights of value 1.

The resulting initial longitudinal baseline weight (s_magzat_v) cannot be used for analytical purposes by itself, but it does serve as a basis for any subsequent cross-sectional or longitudinal weight, whether for the entire population or for a subset of it. In what follows – not only for the 6-month data collection, but also for subsequent data-collection waves – whenever an analysis weight is produced, it will always be used as a starting point, and will be modified in various ways. Naturally, the number of elements for these additional analyses will in future always be lower than 8,844.

6.2. THE 6-MONTH CROSS-SECTIONAL ANALYSIS WEIGHT (S_6HOS)

By designing an initial longitudinal baseline weight, efforts were made to ensure that the main characteristics of the participants in the study were similar to those of the population generally. Of course, unfortunately, during each wave of data collection, we are unable to collect data on everyone who is part of the total cohort population concerned. In the case of the 6-month data collection, we were able to collect data on 8,462 individuals, which means that we have data at the 6-month period for 95.7 per cent of the entire cohort population (so there has been very little drop-out, either temporary or permanent). These 8,462 individuals represent the maximum number of cases available in the 6-month weighted database.

Of course, not all respondents will necessarily be included in each analysis (children raised by a non-biological mother, foreign-language questionnaires, incomplete questionnaires, data replaced by a proxy questionnaire, etc.), and so the variable-level distributions and the item counts of the actual analyses will typically be lower than this theoretical maximum.

Why did not everyone participate in the 6-month study wave?

Some children are known to have been permanently and definitely excluded from the population. A total of 193 such cases were recorded (2.2 per cent): perhaps the child had died, or had been separated from the parent and the new caregiver refused to participate in the research, or the family had emigrated for good (n=44). The most common reason for permanent drop-out, however, was not any of the above, but the fact that some of the mothers categorically refused to respond to the 6-month follow-up.

In another 2.1 per cent of cases, even though for some reason the data were not collected at the 6-month stage, it is possible that data will be collected in the future: for example, if the reason for non-collection was that the family was temporarily based abroad or its whereabouts at the time of the visit were unknown. Also, if the child was not included in the 6-month data collection due to illness or a change in the health visitor, it is almost certain that he or she will participate in the future. (In fact, at the time of writing, several are already participating in the 18-month data-collection wave.)

Obviously, the 95.7 per cent who participated in the 6-month data collection may differ in some respects from the 4.3 per cent who did not. It can also be assumed that participation/drop-out is not entirely random: certain groups from the total population concerned that would otherwise have participated may have been more likely to be missing at the 6-month stage: very young parents, those with at most 8 years of education and unmarried women are slightly more likely to have dropped out of the 6-month data collection. However, given the very low drop-out rate (less than 8 per cent in all the social and demographic groups used for weighting), it can be assumed that the distribution of data on the 8,462 children in the 6-month analysis population will not differ greatly from the 8,844 children in the total population.

To compensate for these (albeit small) biases, we examined the drop-out probability per pooled weight cell and adjusted the initial longitudinal baseline weights of the 6-month population, so that the marginal distributions of the 8,462 individuals in the 6-month analysis population were the same as for the full population by weighting dimension. In this way, we were able to guarantee that the characteristics of the 8,621 individuals matched the characteristics of the entire cohort population. The non-participants in the 6-month survey were assigned a weight of zero, as appropriate. This weight guarantees that the data obtained are 'representative' of the cohort study population in terms of the distributions given. It is thus used in most cross-sectional analyses based on the 6-month data collection, including analyses that include sample members for whom traditional (non-proxy) data collection was also conducted at the prenatal wave.

6.3. MATERNAL WEIGHTING AT THE 6-MONTH STAGE (S_6HOS_A)

In contrast to the data collection during pregnancy (when the pregnant woman seemed a 'reasonable' unit of analysis), at the 6-month stage the basic unit of analysis, research and statistical reporting is already, naturally enough, in most cases the 6-month-old child whose life course will be followed throughout the research. Research questions typically relate to the child. For example, the proportion of children who are breastfed, the risk to the child's development of being born premature, the proportion of children born into a family with a grandparent, etc. However, there are also research questions in the 6-month study wave where the birth mother seems a reasonable and logical unit of analysis: for example, when looking at the impact of the birth experience on a mother's future childbearing plans, or when examining the risk of premature birth. In these cases, the basic unit of analysis and weighting should be the birth mother, rather than the child. For such analyses we have created the '6-month birth mother' weighting (s_6hos_a). The 6-month birth mother database differs from the cross-sectional child weighting for children in two respects. One has to do with twin births: in these cases there are two or three children per mother. The other has to do with children who were not being raised by their birth mother at the age of 6 months (e.g. due to adoption, maternal death or employment abroad). Both of these phenomena are relatively rare: 126 twin births (including one set of triplets) and 29 cases of children who were not being raised by their birth mother at the age of 6 months were reported. Thus, for the 6-month data collection, we have data on a total of 8,311 birth mothers.

The 6-month birth mother weighting for this population of 8,311 was constructed very simply by adjusting the 6-month cross-sectional weight, so that the weighted population of 8,311 birth mothers was equal to the number of participants. Consequently, the resulting maternal and child weighting variables take quite similar values.

The prenatal wave of data collection (typically undertaken 2 months before the child's birth), the population data at the time of labour/birth and the 6-month data collection cover a very short time span compared to traditional follow-up studies – typically 8 months in total. This, and the low drop-out rate between the two data collections, allows us to use the 6-month cross-sectional weights in longitudinal analyses that look at changes that occur between the prenatal wave of data collection and the 6-month data collection, using either the ‘cross-sectional analysis weight’ or the ‘birth mother weight’, depending on the type of analysis. Naturally, for the next data-collection waves, weights of a specific longitudinal nature will have to be developed and used.

7. PSYCHOLOGICAL SCALES

Some of the psychological variables that were the focus of the Cohort ’18 study – maternal generalized anxiety and depression, maternal-infant attachment, parents’ relationship interactions, child temperament – were measured not by individual questions, but with standard scales comprising several questions or statements. The use of multi-item measures provides insights into abstract phenomena that cannot be directly measured using only a single question. Some of the scales are the same as were used in the prenatal wave of the Cohort ’18 study (see Szabó et al., 2021, Appendix 8.3. for their use during pregnancy).

In this section, we first present a preliminary test of the abbreviated versions of two measurement tools. We then provide a detailed review of the multi-item scales used in the 6-month wave of the Cohort ’18 study. In addition to describing the purpose, source, basic properties of the instruments and the variables that are computed based on them, we present descriptive data and an assessment of the internal consistency[1] of the questionnaires obtained from analysis of the 6-month database, as well as their completion rates.

7.1. PRELIMINARY TESTING OF SHORTENED SCALES

At the 6-month wave of the Cohort ’18 study, we planned to measure two constructs – maternal-infant attachment and infant temperament – for which there was no available Hungarian-adapted scale that was short enough for the multidisciplinary data collection (planned to last only half an hour). Therefore, for the maternal-infant attachment, we aimed to test a shortened version of the *Maternal Postnatal Attachment Scale* (MPAS; Condon and Corkindale, 1998); and for infant temperament, we aimed to shorten the *Infant Behavior Questionnaire - Revised - Very Short Form* (IBQ-R-VSF; Putnam et al., 2014).

This subsection presents the preliminary testing of the shortened scales, based mainly on the results of an online questionnaire study. We considered the comprehensibility of the items, the reliability of the abbreviated scales, their correlations with the complete scales, the descriptive data of their total scores and, for the IBQ-R-VSF, the results of a previous large-scale factor analysis (Peterson et al., 2017) and the factor analyses conducted in the present study, as well as the proportion of ‘not applicable’ responses.

To test convergent validity, we analysed the associations of the scales with each other and with maternal depression. Indeed, previous studies have shown that the degree of maternal-infant attachment is inversely associated with maternal depression, as well as

---

[1] Internal consistency is an important indicator of reliability, i.e. whether the items of a given scale capture the same phenomenon. One indicator is the Cronbach’s alpha value, for the interpretation of which there are several guidelines in the literature. Here, we rely on the interpretation by George and Mallery (2003, p. 231), which states that the reliability of the test is unacceptable below 0.5, poor between 0.5 and 0.6, questionable between 0.6 and 0.7, acceptable between 0.7 and 0.8, good between 0.8 and 0.9, and excellent above 0.9.

with a child's difficult temperament (Condon and Corkindale, 1998; Scopesi et al., 2004). Moreover, maternal depressive symptoms may be slightly associated with an infant's difficult temperament (Britton, 2011).

Method

Sample and procedure

The members of groups on a social media site for pregnant women and mothers with young children were asked to take part in the study, specifically those who had a 6-month-old child. Following informed consent, the participants completed an anonymous online questionnaire. The questionnaire consisted of demographic questions and self-reported psychological scales. Data analysis was performed in IBM SPSS Statistics 22.

A total of 166 mothers with children aged 5–9 months completed the questionnaire. Their average age was 31.06 years (SD=4.64). One (0.6 per cent) had at most primary education; 60 (36.2 per cent) had secondary education; and 105 (63.3 per cent) had tertiary education. In terms of their subjective assessment of their socioeconomic status, two (1.2 per cent) considered themselves to belong to the working class; 26 (15.7 per cent) to the lower middle class; 107 (65.5 per cent) to the middle class; and 31 (18.7 per cent) to the upper middle class. Of the participants, 113 (68.1 per cent) had one child; 44 (26.5 per cent) had two children; and nine (5.4 per cent) had three or more children. In all, 119 (71.7 per cent) were married; 43 (25.9 per cent) lived with a partner; and four (2.4 per cent) did not live with a partner.

As for the respondents' children, 79 were boys (47.6 per cent) and 87 were girls (52.4 per cent). In terms of age, they were all between 20 and 36 weeks (M=26.93 weeks, SD=3.86 weeks). Some 154 (92.8 per cent) had been born on time and at appropriate weight; two (1.2 per cent) had had low birth weight; three (1.8 per cent) had been preterm; and seven (4.2 per cent) had been both low birth weight and preterm. According to their mothers, the health of two children (1.2 per cent) was poor; of six (3.6 per cent) it was fair; of 47 (28.3 per cent) it was good; and of 111 children (66.9 per cent) it was excellent.

Measurements

At the start of the online questionnaire, information was collected on certain characteristics of the mothers (age, highest level of education, subjective socioeconomic status, number of children, relationship status) and their children (gender, date of birth, birth weight and age, overall health).

The MPAS questionnaire (Condon and Corkindale, 1998; Hungarian version: Sz. Makó et al., n.d.), which measures maternal-child attachment during the first year of life, contains 19 items with a variable number (between two and five) and content of response categories. Eight items of the questionnaire are reversed. The items are grouped into three sub-scales: quality of attachment (nine items); absence of hostility (five items); and pleasure in interaction (five items). When conducting an evaluation, in order to ensure equal weighting of the questions, it is necessary to recode the responses so that they take a value of between 1 and 5 (Condon, 2015). Six items of the questionnaire were used in the Millennium Cohort Study (Johnson et al., 2012) and in Growing Up in Scotland (Parkes and Wight, 2011), including two statements from each of the three sub-scales (items 1, 5, 9, 11, 15, 19, of which 9 and 11 are reversed). The present study aimed to test this six-item version of the MPAS scale.

The IBQ-R-VSF questionnaire (Putnam et al., 2014; Hungarian version: Lakatos et al., 2014), which measures the temperament of children aged 3–12 months, contains 37 items, including one reversed item. Mothers indicate on a seven-point scale (1 – Never, 7 – Always) how often their child has exhibited certain behaviours in the previous week. In addition, they may also indicate a ‘not applicable’ response category if the child was not observed in the situation described in the question during the previous week. When aggregating the scores of the items, it is necessary to calculate the average of the scores of the items answered: the score of a sub-scale may thus vary from 1 to 7. The items of the questionnaire are grouped into three factors according to Putnam et al. (2014): positive affectivity/surgency (PAS, 13 items); negative emotionality (NEG, 12 items); and orienting and regulatory capacity (ORC, 12 items). The aim of this study was to produce a 15-item version of the IBQ-R-VSF scale, to include five items from each of the three broad factors described by Putnam et al. (2014).

Respondents were also asked how well they understood the statements in the MPAS and IBQ-R-VSF questionnaires. Half (n=83) of the respondents answered this question for one questionnaire and half (n=83) for the other, rating each item on a three-point scale (1 – Very difficult to understand, 2 – Somewhat difficult to understand, 3 – Easy to understand).

Finally, maternal symptoms of depression were measured using the eight-item version of the *Center for Epidemiologic Studies – Depression Scale* (Radloff, 1977) (CES-D-8; Bracke et al., 2008), based on the Hungarian translation by Szeifert (2010). The CES-D-8 scale consists of eight items, on which mothers indicated on a four-point Likert scale how often they had experienced certain symptoms in the previous week (0 – None or almost none of the time (i.e. for less than a day), 3 – All or almost all of the time (i.e. for 5–7 days)). The total score thus varies between 0 and 24 (Cronbach’s alpha=0.772).

Results

Maternal postnatal attachment (MPAS)

Overall, the items of the questionnaire were found to be well understood along the three-point scale applied (M=2.87, SD=0.26), with only 1–3 mothers (1.2–3.6 per cent) finding any given item very difficult to understand and 1–21 mothers (1.2–25.3 per cent) finding it somewhat difficult to understand. The six-item version was also highly understandable (M=2.90, SD=0.28), with item 1 being the least comprehensible (considered somewhat or very difficult to understand by 10.8 per cent of mothers).

The complete scale had good internal consistency (Cronbach’s alpha=0.807), while the six-item version appeared to be acceptable (Cronbach’s alpha=0.595), considering that the six items included questions from three different sub-scales (the six-item scale was found to be less reliable in the Scottish cohort study, with Cronbach’s alpha=0.52 (Parkes and Wight, 2011)).

The overall scores for both the full questionnaire and the shortened version showed a skewed distribution, with a shift towards the high end (see Table 7 for descriptive data). When examined using Spearman’s rank correlation, the total score of the six-item version was highly correlated with the total score of the complete scale (rho=0.844, p<0.001).

Table 7

Descriptive data for the MPAS questionnaire and for its abbreviated version

Scale version	M	SD	Possible range	Sample range	Skewness	Kurtosis	Shapiro–Wilk
Complete –MPAS	81.20	4.42	19-95	49.5–95	−0.964	1.757	p<0.001
Six items	24.78	3.08	6-30	15.9–30	−0.365	−0.140	p=0.006

As for the convergent validity of the shortened questionnaire, it showed similar correlations with the different variables as the complete scale: both were moderately negatively correlated with the maternal depression level (rho=-0.538 for the complete scale, -0.454 for the abbreviated; p<0.001 for both). They were weakly positively correlated with the child's orienting and regulatory capacity (rho=0.354 for the complete scale, 0.379 for the abbreviated; p<0.001 for both). Negative emotionality, however, was more strongly associated only with the full questionnaire (rho=-0.279, p<0.001 for the complete scale; rho=-0.153, p=0.049 for the abbreviated), while positive affectivity/surgency was more strongly associated only with the abbreviated version (rho=0.144, p=0.065 for the complete scale; rho=0.159, p=0.040 for the abbreviated).

Overall, therefore, the six-item version of the questionnaire showed itself a promising tool.

Temperament (IBQ-R-VSF)

To abbreviate the IBQ-R-VSF temperament questionnaire, we relied primarily on the results of an exploratory factor analysis conducted by Peterson et al. (2017), who analysed data on 5,639 children aged 23–52 weeks in New Zealand. They proposed a five-factor structure with the factors: positive affectivity/surgency; negative emotionality; orienting capacity; affiliation/regulation; and fear. Thus, the latter two additional factors were separated out from two of the broader factors (orienting and regulatory capacity and negative emotionality, respectively). Nevertheless, given the small number of items in the proposed abbreviated scale, in this study we focused on the three broad factors (PAS, NEG, ORC) proposed by Putnam et al. (2014) and selected items to represent them. In selecting the abbreviated item sets, we considered: (1) the factor loadings of the item on its corresponding factor, as shown in Peterson et al.'s (2017) large-sample study; (2) the factor loadings obtained as a result of our online survey; (3) the comprehensibility of each item; (4) the proportion of 'not applicable' responses to each item; and (5) the need to ensure that the items chosen encompassed as many dimensions of the original IBQ-R questionnaire as possible. (These data are illustrated later in Table 9.)

The negative emotionality factor contains statements from three sub-dimensions of the IBQ-R questionnaire: sadness; distress to limitations; and fear. To represent this factor, items 4, 16, 23, 29 and 32 were selected from the IBQ-R-VSF questionnaire. The average comprehensibility of the items in the abbreviated questionnaire (M=2.97, SD=0.10) was similar to the comprehensibility of the complete negative emotionality sub-scale (M=2.93, SD=0.16). The percentage of 'not applicable' responses to the selected items ranged from 2.4 per cent to 12 per cent. The abbreviated set of items was highly correlated with the complete sub-scale (rho=0.913, p<0.001). The internal consistency of the mean score for these items was acceptable (Cronbach's alpha=0.611). Reliability was higher among children aged 5.5 to 7 months (n=102), who were closer in age to the target group of the 6-month wave of the Cohort '18 study (Cronbach's alpha=0.646).

The positive affectivity/surgency factor contains statements from six dimensions of the IBQ-R questionnaire: activity level; smiling and laughter; high intensity pleasure; vocal reactivity; approach; and perceptual sensitivity. To represent this factor, items 1, 13, 14, 15 and 21 were selected from the IBQ-R-VSF questionnaire. The average comprehensibility of the items in the abbreviated questionnaire (M=2.97, SD=0.10) was similar to the comprehensibility of the complete positive affectivity/surgency sub-scale (M=2.94, SD=0.13). The percentage of 'not applicable' responses for the selected items ranged from 0 per cent to 24.1 per cent. The abbreviated set of items was moderately correlated with the complete sub-scale (rho=0.633, p<0.001). The internal consistency of the mean score for these items was acceptable (Cronbach's alpha=0.509), considering the diversity of the construct. Among children aged 5.5 to 7 months (n=102), closer in age to the target group of the 6-month wave of the Cohort '18 study, reliability was higher (Cronbach's alpha=0.591).

The orienting and regulatory capacity factor contains statements from four dimensions of the IBQ-R questionnaire: duration of orienting; low intensity pleasure; cuddliness; and soothability. To represent this factor, items 24, 25, 30, 31 and 34 were selected from the IBQ-R-VSF questionnaire. The average comprehensibility of the items in the abbreviated questionnaire (M=2.97, SD=0.13) was similar to the comprehensibility of the complete orienting and regulatory capacity sub-scale (M=2.95, SD=0.12). The percentage of 'not applicable' responses for the selected items ranged from 0 per cent to 12 per cent. The abbreviated set of items was highly correlated with the complete sub-scale (rho=0.844, p<0.001). The internal consistency of the mean score for these items was on the borderline of acceptability (Cronbach's alpha=0.500), taking into account the diversity of the construct. Among children aged 5.5 to 7 months (n=102), closer in age to the target group of the 6-month wave of the Cohort '18 study, reliability was higher (Cronbach's alpha=0.525).

The overall scores on both the complete and the shortened sub-scales differed significantly from the normal distribution (see Table 8 for descriptive data). The mean scores for the abbreviated sub-scales were slightly higher than for the complete versions.

Table 8

Descriptive data for the sub-scales of the IBQ-R-VSF questionnaire and for their abbreviated versions

Scale version	M	SD	Possible range	Sample range	Skewness	Kurtosis	Shapiro–Wilk
Complete NEG	3.79	1.09	1–7	1.55–6.27	0.076	–0.804	p=0.033
Abbreviated NEG	4.16	1.29	1–7	1–6.40	–0.169	–0.964	p=0.001
Complete PAS	5.18	0.81	1–7	2.38–6.62	–0.687	0.564	p=0.001
Abbreviated PAS	6.01	0.79	1–7	2.40–7	–1.196	2.225	p<0.001
Complete ORC	5.66	0.69	1–7	2.58–7	–0.556	1.457	p=0.004
Abbreviated ORC	5.97	0.78	1–7	2.40–7	–1.260	2.587	p<0.001

IBQ-R-VSF – Infant Behavior Questionnaire - Revised - Very Short Form, NEG – negative emotionality sub-scale, PAS – positive affectivity/surgency sub-scale, ORC – orienting and regulatory capacity sub-scale.

As for the convergent validity of the abbreviated question sets, they showed correlations similar to the complete versions. Negative emotionality was weakly and positively correlated with the maternal depression level (rho=0.244 for the complete sub-scale, 0.217 for the abbreviated; p<0.01 for both), while orienting and regulatory capacity was negatively correlated (rho=-0.338 for the complete sub-scale, -0.298 for the abbreviated; p<0.001 for both). Maternal depression and positive affectivity/surgency, however, were only related with regards to the abbreviated item set (rho=-0.186, p=0.017), not for the complete sub-scale (rho=-0.113, p=0.151). For maternal-infant attachment, a similar phenomenon was observed: negative emotionality was negatively related (rho=-0.227, p=0.003), while orienting and regulatory capacity (rho=0.322, p<0.001) and positive affectivity/surgency (rho=0.195, p=0.012) were positively related to the shortened item set, which differed from the full sub-scale for positive affectivity/surgency.

Overall, based on the factor analysis by Peterson et al. (2017) and the present online survey, the selected items of the abbreviated sub-scales adequately represent the three main dimensions of the survey; moreover, they have proved to be well-understood and answerable questions with a high response rate in a Hungarian sample. However, the internal consistency of the sub-scales is rather low, which is due to the small number of items measuring multiple dimensions.

Table 9

Items of IBQ-R-VSF; their classification into IBQ-R sub-factors and IBQ-R-VSF broad factors, factor loadings, applicability and comprehensibility

Item	IBQ-R sub-factor	IBQ-R-VSF factor	Factor loading (Peterson et al., 2017)[a]	Items	Factor loading[b]	NA %	Diff %
33	Fear	NEG	Fear: 0.832	When in the presence of several unfamiliar adults, how often did the baby cling to a parent?	0.412	19.9	6
4	**Fear**	**NEG**	**Fear: 0.793**	**When introduced to an unfamiliar adult, how often did the baby cling to a parent?**	**0.371**	**12**	**4.8**
28	Fear	NEG	Fear: 0.769	When introduced to an unfamiliar adult, how often did the baby refuse to go to the unfamiliar person?	0.405	22.9	9.6
17	Fear	NEG	NEG: 0.390	How often during the last week did the baby startle at a sudden change in body position (e.g. when moved suddenly)?	0.412	7.2	6
16	**Distress to Limitations**	**NEG**	**NEG: 0.682**	**How often did the baby seem angry (crying and fussing) when you left her/him in the crib?**	**0.657**	**9.6**	**4.8**
10	Distress to Limitations	NEG	NEG: 0.634	After sleeping, how often did the baby cry if someone doesn't come within a few minutes?	0.589	9	6
32	**Distress to Limitations**	**NEG**	**NEG: 0.615**	**When the baby wanted something, how often did s/he become upset when s/he could not get what s/he wanted?**	**0.631**	**4.8**	**3.6**
23	**Distress to Limitations**	**NEG**	**NEG: 0.443**	**How often during the last week did the baby protest being placed in a confining place (infant seat, play pen, car seat, etc.)?**	**0.601**	**2.4**	**1.2**
29	**Sadness**	**NEG**	**NEG: 0.643**	**When you were busy with another activity, and your baby was not able to get your attention, how often did s/he cry?**	**0.692**	**4.2**	**1.2**
9	Sadness	NEG	NEG: 0.531	When it was time for bed or a nap and your baby did not want to go, how often did s/he whimper or sob?	0.463	4.8	6
3	Sadness	NEG	NEG: 0.500	When tired, how often did your baby show distress?	0.369	1.8	13.3
22	Sadness	NEG	NEG: 0.440	At the end of an exciting day, how often did your baby become tearful?	0.580	6	6

[a] The reported factor loadings show the results of Peterson et al.'s (2017) five-factor exploratory factor analysis (Maximum Likelihood and oblique rotation).

[b] The reported factor loadings show the results obtained by examining the given sub-scales separately using single factor exploratory factor analyses (Maximum Likelihood) based on our online survey.

Items selected for use in the Cohort '18 study are in bold.

IBQ-R – Infant Behavior Questionnaire - Revised, IBQ-R-VSF – Infant Behavior Questionnaire - Revised - Very Short Form, NA % – percentage of answers not applicable, based on our online survey, Diff % – percentage of answers very difficult and somewhat difficult to understand, based on our online survey, NEG – negative emotionality factor, PAS – positive affectivity/surgency factor, ORC – orienting and regulatory capacity factor, AFF – affiliation/regulation factor, OC – orienting capacity factor.

Table 9

Items of IBQ-R-VSF; their classification into IBQ-R sub-factors and IBQ-R-VSF broad factors, factor loadings, applicability and comprehensibility (continued)

Item	IBQ-R sub-factor	IBQ-R-VSF factor	Factor loading (Peterson et al., 2017)[a]	Items	Factor loading[b]	NA %	Diff %
13	**Activity Level**	**PAS**	**PAS: 0.502**	**When placed on his/her back, how often did the baby squirm and/or turn body?**	**0.450**	**2.4**	**2.4**
1	**Activity Level**	**PAS**	**PAS: 0.477**	**When being dressed or undressed during the last week, how often did the baby squirm and/or try to roll away?**	**0.381**	**0**	**3.6**
37	Activity Level	PAS	PAS: 0.245	When placed in an infant seat or car seat, how often did the baby squirm and turn body?	0.239	1.8	3.6
14	**High Intensity Pleasure**	**PAS**	**PAS: 0.499**	**During a peekaboo game, how often did the baby laugh?**	**0.545**	**6**	**1.2**
2	High Intensity Pleasure	PAS	PAS: 0.487	When tossed around playfully how often did the baby laugh?	0.483	13.9	4.8
7	Approach	PAS	PAS: 0.480	How often during the week did your baby move quickly toward new objects?	0.441	15. 1	12
20	Approach	PAS	PAS: 0.385	When visiting a new place, how often did your baby get excited about exploring new surroundings?	0.453	18.7	6
21	**Smiling and Laughter**	**PAS**	**PAS: 0.402**	**How often during the last week did the baby smile or laugh when given a toy?**	**0.506**	**0.6**	**3.6**
8	Smiling and Laughter	PAS	PAS: 0.394	When put into the bath water, how often did the baby laugh?	0.480	2.4	4.8
15	**Perceptual Sensitivity**	**PAS**	**PAS: 0.371**	**How often does the infant look up from playing when the telephone rings?**	**0.493**	**24.1**	**3.6**
27	Perceptual Sensitivity	PAS	PAS: 0.335	How often did your baby notice the sound of an airplane passing overhead?	0.441	56.6	9.6
36	Vocal Reactivity	PAS	PAS: 0.347	How often did your baby make talking sounds when riding in a car?	0.366	16.9	8.4
26	Vocal Reactivity	PAS	PAS: 0.239	When hair was washed, how often did the baby vocalize?	0.248	4.8	9.6

[a] The reported factor loadings show the results of Peterson et al.'s (2017) five-factor exploratory factor analysis (Maximum Likelihood and oblique rotation).

[b] The reported factor loadings show the results obtained by examining the given sub-scales separately using single factor exploratory factor analyses (Maximum Likelihood) based on our online survey.

Items selected for use in the Cohort '18 study are in bold.

IBQ-R – Infant Behavior Questionnaire - Revised, IBQ-R-VSF – Infant Behavior Questionnaire - Revised - Very Short Form, NA % – percentage of answers not applicable, based on our online survey, Diff % – percentage of answers very difficult and somewhat difficult to understand, based on our online survey, NEG – negative emotionality factor, PAS – positive affectivity/surgency factor, ORC – orienting and regulatory capacity factor, AFF – affiliation/regulation factor, OC – orienting capacity factor.

Table 9

Items of IBQ-R-VSF; their classification into IBQ-R sub-factors and IBQ-R-VSF broad factors, factor loadings, applicability and comprehensibility (continued)

Item	IBQ-R sub-factor	IBQ-R-VSF factor	Factor loading (Peterson et al., 2017)[a]	Items	Factor loading[b]	NA %	Diff %
34	**Cuddliness**	**ORC**	**AFF: 0.712**	**When rocked or hugged, in the last week, did your baby seem to enjoy him/herself?**	**0.602**	**7.8**	**2.4**
24	**Cuddliness**	**ORC**	**AFF: 0.521**	**When being held, in the last week, did your baby seem to enjoy him/herself?**	**0.438**	**0**	**2.4**
11	Cuddliness	ORC	PAS: 0.238	In the last week, while being fed in your lap, how often did the baby seem eager to get away as soon as the feeding was over?	0.101	20.5	9.6
30	**Low Intensity Pleasure**	**ORC**	**AFF: 0.607**	**How often during the last week did the baby enjoy gentle rhythmic activities, such as rocking or swaying?**	**0.569**	**1.8**	**2.4**
5	Low Intensity Pleasure	ORC	OC: 0.346	How often during the last week did the baby enjoy being read to?	0.533	26.5	7.2
18	Low Intensity Pleasure	ORC	OC: 0.286	How often during the last week did the baby enjoy hearing the sound of words, as in nursery rhymes?	0.520	10.2	3.6
35	Soothability	ORC	AFF: 0.418	When patting or gently rubbing some part of the baby's body, how often did s/he soothe immediately?	0.563	7.8	7.2
25	**Soothability**	**ORC**	**AFF: 0.340**	**When showing the baby something to look at, how often did s/he soothe immediately?**	**0.531**	**2.4**	**4.8**
12	Soothability	ORC	AFF: 0.302	When singing or talking to your baby, how often did s/he soothe immediately?	0.422	2.4	1.2
19	Duration of Orienting	ORC	OC: 0.632	How often during the last week did the baby look at pictures in books and/or magazines for 5 minutes or longer at a time?	0.509	20.5	4.8
31	**Duration of Orienting**	**ORC**	**OC: 0.562**	**How often during the last week did the baby stare at a mobile, crib bumper or picture for 5 minutes or longer?**	**0.562**	**12**	**8.4**
6	Duration of Orienting	ORC	OC: 0.486	How often during the last week did the baby play with one toy or object for 5–10 minutes?	0.281	1.8	6

[a] The reported factor loadings show the results of Peterson et al.'s (2017) five-factor exploratory factor analysis (Maximum Likelihood and oblique rotation).

[b] The reported factor loadings show the results obtained by examining the given sub-scales separately using single factor exploratory factor analyses (Maximum Likelihood) based on our online survey.

Items selected for use in the Cohort '18 study are in bold.

IBQ-R – Infant Behavior Questionnaire - Revised, IBQ-R-VSF – Infant Behavior Questionnaire - Revised - Very Short Form, NA % – percentage of answers not applicable, based on our online survey, Diff % – percentage of answers very difficult and somewhat difficult to understand, based on our online survey, NEG – negative emotionality factor, PAS – positive affectivity/surgency factor, ORC – orienting and regulatory capacity factor, AFF – affiliation/regulation factor, OC – orienting capacity factor.

7.2. CHARACTERISTICS OF THE SCALES AND THE COMPUTED VARIABLES

The majority of the scales used in the Cohort '18 6-month study wave, with the exception of the temperament questionnaire, are contained in the self-administered questionnaire, which only biological mothers were asked to complete (n=8241). The full dataset for this questionnaire is missing for 3.8 per cent of the mothers who responded in the 6-month wave (n=312). Either they did not complete it at all, or possibly their questionnaire could not be linked to their additional data due to a lack of correct identification. (This subsection discusses the completion of the multi-item scales in the self-administered questionnaire; information on the completion of the other questionnaire modules is provided in section 5.)

Generalized anxiety

To assess generalized anxiety symptoms, the Hungarian translation of the same *Generalized Anxiety Disorder-2* scale (GAD-2; Kroenke et al., 2007) was used as in the prenatal wave (available to download at http://www.phqscreeners.com). This ultra-brief scale is a quick screening tool for anxiety disorders: above a certain threshold, the incidence of an anxiety disorder is likely. When used as a continuous variable, a higher total score indicates a higher prevalence of generalized anxiety symptoms.

Respondents indicated on a four-point Likert scale how often they had experienced certain feelings or behaviours in the previous two weeks (1 – Not at all, 4 – Nearly every day). Neither of the two items was reversed. The questionnaire was originally designed to be answered on a scale of 0–3, but in order to maintain a consistent structure and easier response in the Cohort '18 questionnaire the scores are 1–4. Therefore, when preparing the database the values in the Cohort '18 were recoded for further analysis. The data for descriptive statistics with these recoded 0–3 values, consistent with the literature, are presented in this methodological volume. (These recoded values are also included in the public database.) When evaluation is undertaken, the items are summed to form a total score, which can vary between 0 and 6.

Table 10

Measuring generalized anxiety in the 6-month wave of the Cohort '18 study

Measured phenomenon	Generalized anxiety
Sample	n=8241 mothers
Scale	Generalized Anxiety Disorder-2
Number of items used	2
Cronbach's alpha	0.610
Spearman–Brown coefficient	0.619
Mean ± standard deviation	1.04 ± 1.16
Median	1
Recoding	1→0, 2→1, 3→2, 4→3
Fully completed response rate	87.4%
Missing data replacement	Did not apply
Computed variable	s2mgads
Variables used	s2mgad1 – s2mgad2

The questionnaire was completed in full by 87.4 per cent of mothers. The internal consistency of its total score, which was also checked with the Spearman–Brown coefficient due to its two-item nature, was found to be questionable. Given that it consists of only two items, no data replacement was applied. Descriptive data for the scale are presented in Table 10.

Depression

Symptoms of depression were assessed in the same way as during the prenatal wave, using the eight-item version of the *Center for Epidemiologic Studies – Depression Scale* (Radloff, 1977) (CES-D-8; Bracke et al., 2008), based on the Hungarian translation by Szeifert (2010). Respondents indicated on a four-point Likert scale how often they had experienced certain feelings or behaviours in the previous week (1 – None or almost none of the time (for less than 1 day), 4 – All or almost all of the time (for 5–7 days)). Items 4 and 6 of the questionnaire are reversed. The questionnaire was originally designed to be answered on a scale of 0–3, and so the values of 1–4 used in the Cohort '18 questionnaire were recoded when the database was prepared, and the descriptive statistics with the recoded values of 0–3 are presented in this methodological volume. (The public database also contains these recoded values.) When evaluation is undertaken, items are summed to form a total score, which can vary between 0 and 24.

The questionnaire was completed fully by 86.9 per cent of respondents. Given that many participants missed out only one item, questionnaire data above 80 per cent completion were replaced with averages for the respondent's own completed responses in order to generate the total score, a promising method for item-level missing data replacement (Bono et al., 2007; Downey and King, 1998). Some 91.1 per cent of participants answered at least 80 per cent of the questionnaire (six items). The reliability of the scale was found to be acceptable both before and after data replacement. Descriptive data for the scale are presented in Table 11.

Table 11

Measuring depression in the 6-month wave of the Cohort '18 study

Measured phenomenon	Depression
Sample	n=8241 mothers
Scale	Center for Epidemiologic Studies – Depression
Number of items used	8
Cronbach's alpha (raw; after data replacement)	0.773; 0.776
Mean ± standard deviation (raw; after data replacement)	3.81 ± 3.45; 3.86 ± 3.49
Median (raw; after data replacement)	3; 3
Recoding	1→0, 2→1, 3→2, 4→3
Fully completed response rate	86.9%
Missing data replacement	In case of at least six items answered
Response rate after data replacement	91.1%
Computed variable	s2mdeprs
Variables used	s2mdepr1 – s2mdepr8

Maternal-child attachment

Maternal attachment to the infant was measured using six items selected from the Hungarian translation (Sz. Makó et al., n.d.) of the *Maternal Postnatal Attachment Scale* (MPAS; Condon and Corkindale, 1998), and was based on the abbreviated version used in the Millennium Cohort Study (Johnson et al., 2012) and in Growing Up in Scotland (Parkes and Wight, 2011). The abbreviated item set contains two questions from each of the original three factors (quality of attachment, absence of hostility and pleasure in interaction).

The number of response options for the questions varies between four and five. Items 3 and 4 of the scale are reversed. In order to ensure equal weighting of the questions, it is necessary to rescore the responses to take a value between 1 and 5 (Condon, 2015). This allows the total score of the questionnaire to have a value of between 6 and 30.

The questionnaire was completed in full by 92.3 per cent of respondents. The reliability of the scale was found to be poor, and it would not be increased by the deletion of any of the items. However, the value is acceptable considering that the six items cover three dimensions of maternal-child attachment. Descriptive data for the scale are presented in Table 12.

Table 12

Measuring maternal-child attachment in the 6-month wave of the Cohort '18 study

Measured phenomenon	Maternal-child attachment
Sample	n=8241 mothers
Scale	Maternal Postnatal Attachment Scale
Number of items used	6
Cronbach's alpha	0.564
Mean ± standard deviation	26.31 ± 2.85
Median	26.6
Recoding	For s2mpa2, s2mpa5 and s2mpa6: 1→1, 2→2.3, 3→3.6, 4→5
Fully completed response rate	92.3%
Missing data replacement	Did not apply
Computed variable	s2mpas
Variables used	s2mpa1 – s2mpa6

Relationship interactions

Regarding the quality of the parents' relationship, the frequency of positive and negative interactions in their relationship was measured using the Hungarian version (Kopcsó, 2018) of the 11-item *Gilford–Bengtson Marital Satisfaction Scale* (Gilford and Bengtson, 1979; Silverstein and Bengtson, 2008), which was also used in the prenatal wave of the study.

Respondents indicated on a five-point Likert scale how often they thought the given couple interactions occurred (1 – Hardly ever, 5 – Always). None of the items were reversed. For evaluation, a total score per sub-scale is required, ranging from 1 to 25 for positive interactions and 1 to 30 for negative interactions.

At the 6-month wave, the positive interaction sub-scale was completed in full by 88.9 per cent of respondents, and 80 per cent of the questions (at least four items) were answered by 91 per cent of participants. Regarding the negative interaction sub-scale, 87.4 per cent of respondents completed the negative interaction sub-scale and 90.7 per cent answered at least 80 per cent of the questions (at least five items). The reliability of the relationship conflict sub-scale was found to be acceptable, while the reliability of the positive interaction sub-scale was found to be good. Descriptive data for the sub-scales are presented in Table 13.

Table 13

Measuring relationship interactions in the 6-month wave of the Cohort '18 study

Measured phenomenon	Relationship conflicts	Positive relationship interaction
Sample	n=8241 mothers	n=8241 mothers
Scale	Gilford–Bengtson Marital Satisfaction Scale	Gilford–Bengtson Marital Satisfaction Scale
Number of items used	6	5
Cronbach's alpha (raw; after data replacement)	0.756 0.758	0.864 0.865
Mean ± standard deviation (raw; after data replacement)	9.48 ± 3.04 9.51 ± 3.07	20.58 ± 3.84 20.56 ± 3.85
Median (raw; after data replacement)	9 9	21 21
Recoding	Did not apply	Did not apply
Fully completed response rate	87.4%	88.9%
Missing data replacement	In case of at least five items answered	In case of at least four items answered
Response rate after data replacement	90.7%	91%
Computed variable	s2mrnegs	s2mrpozs
Variables used	s2mrneg1 – s2mrneg6	s2mrpoz1 – s2mrpoz5

Temperament

The temperament of infants was assessed using 15 items from the *Infant Behavior Questionnaire - Revised - Very Short Form* (IBQ-R-VSF; Putnam et al., 2014), selected from the Hungarian translation (Lakatos et al., 2014) (see subsection 7.1 for the process of the abbreviation). The applied item sets contain five questions on each of the three broad factors of the IBQ-R-VSF (positive affectivity/surgency, negative emotionality and orienting and regulatory capacity), none of which are reversed.

The temperament scale, unlike previous scales, was part of the main questionnaire that was administered during the personal interview. Thus, the data on this scale were expected for a total of 8,392 children (including 243 twins and the 29 children cared for by a non-birth mother).

Respondents indicated on a seven-point scale (1 – Never, 7 – Always) how often the child had exhibited certain behaviours in the previous week. They could also choose a 'not applicable' response category if the child had not been observed in the situation described in the question during the previous week. During evaluation, 'not applicable' responses were treated as missing responses, and so only questions with a frequency response were considered to have been answered when response rates were calculated.

For evaluation, an aggregate score is calculated as the average of the scores of the items answered, according to the official guidelines: the score of a sub-scale can therefore vary between 1 and 7. For 83.8 per cent of the children, all items were answered; for 95.9

per cent only one question remained unanswered; for 98.8 per cent – up to two; and for 99.5 per cent – up to three questions. On a sub-scale level, the negative emotionality sub-scale scores exceeded the 20 per cent gap (less than four items answered) for 1.3 per cent of children. For the positive affectivity/surgency and orienting and regulatory capacity sub-scales, the rate was 0.4 per cent. In these cases, the overall mean score was not calculated.

The reliability of the positive affectivity/surgency and the orienting and regulatory capacity sub-scales was found to be weak, and deleting either item would not increase the reliability significantly. The internal consistency of the negative emotionality sub-scale was acceptable. Descriptive data for the sub-scales are presented in Table 14.

Table 14

Measuring child temperament in the 6-month wave of the Cohort '18 study

Measured phenomenon	Negative emotionality	Positive affectivity/surgency	Orienting and regulatory capacity
Sample	n=8392 children	n=8392 children	n=8392 children
Scale	Infant Behavior Questionnaire - Revised - Very Short Form	Infant Behavior Questionnaire - Revised - Very Short Form	Infant Behavior Questionnaire - Revised - Very Short Form
Number of items used	5	5	5
Cronbach's alpha	0.716	0.548	0.529
Mean ± standard deviation[a]	4.13 ± 1.32	6.28 ± 0.68	6.27 ± 0.64
Median[a]	4.2	6.4	6.4
Recoding	Did not apply	Did not apply	Did not apply
Fully completed response rate	91.9%	91.5%	96.6%
At least 80% completion rate	98.7%	99.6%	99.6%
Computed variable[a]	q2chtnegs	q2chtpass	q2chtorcs
Variables used	q2chtempstranger q2chtempbed q2chtempbound q2chtempcry q2chtempwant	q2chtempdress q2chtempback q2chtemplough q2chtempphone q2chtempsmile	q2chtemphold q2chtempcalm q2chtemprock q2chtempwatch q2chtempenjoy

[a] The computed variable and its central values were created when at least 80 per cent of the items of the sub-scale were completed.

REFERENCES

Bono, C., Ried, L.D., Kimberlin, C. and Vogel, B. (2007). Missing data on the Center for Epidemiologic Studies Depression Scale: A comparison of 4 imputation techniques. *Research in Social and Administrative Pharmacy*, *3*(1), 1–27. https://doi.org/10.1016/j.sapharm.2006.04.001

Bracke, P., Levecque, K. and Van de Velde, S. (2008). *The psychometric properties of the CES-D 8 depression inventory and the estimation of cross-national differences in the true prevalence of depression*. Ghent: Ghent University.

Britton, J.R. (2011). Infant temperament and maternal anxiety and depressed mood in the early postpartum period. *Women & Health*, *51*(1), 55–71. https://doi.org/10.1080/03630242.2011.540741

Condon, J.T. (2015). Maternal Postnatal Attachment Scale [Measurement Instrument]. https://doi.org/10.25957/5DC0F28D14338

Condon, J.T. and Corkindale, C.J. (1998). The assessment of parent-to-infant attachment: Development of a self-report questionnaire instrument. *Journal of Reproductive and Infant Psychology*, *16*(1), 57–76. https://doi.org/10.1080/02646839808404558

Downey, R.G. and King, C.V. (1998). Missing data in Likert ratings: A comparison of replacement methods. *The Journal of General Psychology*, *125*(2), 175–191. https://doi.org/10.1080/00221309809595542

George, D. and Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Gilford, R. and Bengtson, V. (1979). Measuring marital satisfaction in three generations: Positive and negative dimensions. *Journal of Marriage and Family*, *41*(2), 387–398. https://doi.org/10.2307/351705

Johnson, J., Atkinson, M. and Rosenberg, R. (2012). *Millennium Cohort Study, psychological, developmental and health inventories*. London: Centre for Longitudinal Studies.

Kopcsó, K. (2018). Scale adaptation. In: Zs. Veroszta (ed.). *Technical report. Growing Up in Hungary – Cohort '18. Hungarian birth cohort study. Prenatal research, preparational phase,* Working Papers on Population, Family and Welfare, No. 30, (pp. 24–29). Budapest: HDRI.

Kroenke, K., Spitzer, R.L., Williams, J.B.W., Monahan, P.O. and Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, *146*(5), 317–325.

Lakatos, K., Tóth, I. and Gervai, J. (2014). Csecsemő Viselkedés Kérdőív [Infant behavior questionnaire]. Unpublished manuscript.

Parkes, A. and Wight, D. (2011). *Growing Up in Scotland: Parenting and children's health*. Edinburgh: Scottish Government.

Peterson, E.R., Waldie, K.E., Mohal, J., Reese, E., Carr, P.E.A., Grant, C.C. and Morton, S.M.B. (2017). Infant Behavior Questionnaire–Revised Very Short Form: A new factor structure's associations with parenting perceptions and child language outcomes. *Journal of Personality Assessment*, *99*(6), 561–573. https://doi.org/10.1080/00223891.2017.1287709

Putnam, S.P., Helbig, A.L., Gartstein, M.A., Rothbart, M.K. and Leerkes, E. (2014). Development and assessment of short and very short forms of the Infant Behavior Questionnaire–Revised. *Journal of Personality Assessment*, *96*(4), 445–458. https://doi.org/10.1080/00223891.2013.841171

Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Scopesi, A., Viterbori, P., Sponza, S. and Zucchinetti, P. (2004). Assessing mother-to-infant attachment: The Italian adaptation of a self-report questionnaire. *Journal of Reproductive and Infant Psychology*, *22*(2), 99–109. https://doi.org/10.1080/02646830042000205963

Silverstein, M. and Bengtson, V.L. (2008). *Longitudinal Study of Generations, 1971, 1985, 1988, 1991, 1994, 1997, 2000, 2005 [California]*. ICPSR - Interuniversity Consortium for Political and Social Research. https://doi.org/10.3886/ICPSR22100.v4

Sz. Makó, H., Szentiványi-Makó, N. and Deák, A. (n.d.). MPAS magyar változat [The Hungarian version of MPAS]. Unpublished manuscript.

Szabó, L., Boros, J., Fábián, I., Gresits, G., Hortobágyi, A., Kapitány, B., Kis, R., Kopcsó, K., Leitheiser, F., Rohr, A., Spéder, Zs. and Veroszta, Zs. (2021). *Growing Up in Hungary – Cohort '18 Hungarian Birth Cohort Study. Technical report 2. Prenatal wave.* Working Papers on Population, Family and Welfare, No. 38. Budapest: HDRI.

Szeifert, L. (2010). Depresszió és életminőség krónikus veseelégtelenségben szenvedő betegek körében [Depression and quality of life in patients with chronic kidney disease], doctoral dissertation. Budapest.

LIST OF WORKING PAPERS

1. László Hablicsek, Pál Péter Tóth: The Role of International Migration in Maintaining the Population Size of Hungary between 2000–2050
2. Maritetta Pongrácz: Birth out of Wedlock
3. Attila Melegh: East/West Exclusions and Discourses on Population in the 20th Century
4. Zsolt Spéder: Fertility and Structural Change in Hungary
5. Sándor Illés: Foreigners in Hungary: Migration from the European Union
6. Magdalena Muszyńska: Family Models in Europe in the Context of Women's Status
7. Attila Melegh, Elena Kondratieva, Perttu Salmenhaare, Annika Forsander, László Hablicsek, Adrienn Hegyesi: Globalisation, Ethnicity and Migration. The Comparison of Finland, Hungary and Russia
8. Zsolt Spéder, Balázs Kapitány: Poverty and Deprivation: Assessing Demographic and Social Structural Factors
9. Etelka Daróczi: Ageing and Health in the Transition Countries of Europe – the Case of Hungary
10. Péter Őri: Demographic Patterns and Transitions in 18–20th Century Hungary. County Pest-Pilis-Solt-Kiskun in the Late 18th and Early 20th Centuries
11. Zsolt Spéder, Balázs Kaptány: Ideational Factors and Parenthood. A Gender- and Parity Specific Analysis in a Post-Communist Society
12. Irén Gödri: The Role of Ethnicity and Social Capital in Immigration to Hungary
13. Attila Melegh, Arland Thornton, Dimiter Philipov, Linda Young-DeMarco: Mapping Societal Developmental Hierarchies in Europe: a Bulgarian Perspective
14. Balázs Kapitány, Zsolt Spéder: Factors Affecting the Realisation of Child-Bearing Intentions in Four European Countries
15. Zsolt Spéder, Balázs Kapitány: Realising Birth Intention in European Comparison – Understanding the Post-Communist Fertility Transition
16. Tamás Faragó: Historical Demography in Hungary: a History of Research
17. Attila Melegh: Net Migration and Historical Development in Southeastern Europe since 1950
18. Róbert I. Gál, Endre Szabó, Lili Vargha: The age profile of invisible transfers: the true size of asymmetry in inter-age reallocations
19. Irén Gödri, Béla Soltész, Boróka Bodacz-Nagy: Immigration or emigration country? Migration trends and their socio-economic background in Hungary: A longer-term historical perspective
20. Péter Őri, Levente Pakot: Residence patterns in nineteenth century Hungary: Evidence from the Hungarian MOSAIC sample.
21. Zsuzsa Blaskó: Studying emigration by extending a large-scale household survey. Methodology, evaluation and descriptive findings
22. Lili Vargha, Róbert Iván Gál, Michelle O. Crosby-Nagy: Household production and consumption over the lifecycle: the National Time Transfer Accounts in 14 European countries
23. Kálmán Joubert, Gyula Gyenis: The Hungarian Longitudinal Growth Study: From birth to the age of 18 years
24. Róbert Iván Gál, Árpád Törzsök: The Savings Gap in Hungary
25. Lajos Bálint: Suicide in the Hungarian Kingdom
26. Róbert Iván Gál, Pieter Vanhuysse, Lili Vargha: Pro-elderly welfare states within pro-child societies: Incorporating family cash and time into intergenerational transfers analysis
27. Zsolt Spéder, Lívia Murinkó, Lívia Sz. Oláh: Sailing close to the wind?
    The effects of third birth policies in post-communist Hungary
28. Róbert Iván Gál, Márton Medgyesi: Financing the lifecycle or mitigating poverty: Redistribution in the Hungarian welfare system by age and income
29. Irén Gödri, Gábor Attila Feleky: Selection of migrants and realization of migration intentions – lessons from a panel study
30. Zsuzsanna Veroszta (ed.): Technical Report. Growing Up in Hungary - Cohort '18 Hungarian Birth Cohort Study. Prenatal Research, Preparational phase.

31. Róbert Iván Gál, Márta Radó: Participation and Postponed Retirement in Central and Eastern Europe
32. Zsuzsanna Veroszta (ed.): Conceptual framework. Growing Up in Hungary – Cohort '18 Hungarian birth cohort study
33. Mikołaj Szołtysek, Bartosz Ogórek: Global and local correlations of Hajnal's historical household formation markers: a geographically weighted approach
34. Laura Szabó, Igor Kiss, Branislav Šprocha, Zsolt Spéder: Fertility of Roma minorities in Central and Eastern Europe
35. Róbert Iván Gál, Márton Medgyesi, Pieter Vanhuysse: The transfer cost of parenthood in Europe
36. Csaba G. Tóth: Multi-population models to handle mortality crises in forecasting mortality: a case study from Hungary
37. Laura Szabó: Mixed-ethnic partnerships and ethnic reproduction among Roma women in Hungary
38. Laura Szabó: Growing up in Hungary - Cohort '18 Hungarian Birth Cohort Study. Technical report 2. Prenatal wave
39. Zsuzsanna Veroszta, Julianna Boros, Balázs Kapitány, Krisztina Kopcsó, Fruzsina Leitheiser, Laura Szabó, Zsolt Spéder: Pregnancy in Hungary - Report on the first wave of the Cohort '18 - Growing up in Hungary
40. Zsuzsanna Veroszta, Julianna Boros, Balázs Kapitány, Krisztina Kopcsó, Fruzsina Leitheiser, Nikolett Gabriella Sándor, Laura Szabó, Zsolt Spéder: Infancy in Hungary. Report on the Second Wave of Cohort '18 – Growing Up in Hungary

The above Working Papers can be ordered at the following e-mail address:
workingpapers@demografia.hu

The Working Papers are also available online at:
http://www.demografia.hu/workingpapers