

Participation and weighting of the national survey at age 10½

To support analyses of the data collected by the survey wave at age 10½, participation and weighting variables were calculated.

Participation

Five participation variables were created to characterize the participation of the ELFE families in the CATI and CAPI surveys.

The variables A10X_PARTIMEN10A_CATI and A10X_PARTIMEN10A_CAPI simply characterize the household's participation in the CATI (telephon survey) and CAPI (home survey) respectively.

A10X_PARTIMEN10A_CATI	Frequency	Percentage
0.0=the household is not part of the surveyable frame (not to be investigate)	5,242	28.6
1.0=the household did not participate (CATI) AND did not respond to the minimum Q	2,622	14.31
1.1=the household did not participate (CATI) - minimum Q completed	658	3.59
2.0=the household participated (CATI started)	9,807	53.51

A10X_PARTIMEN10A_CAPI	Frequency	Percentage
0.0=the household is not part of the surveyable frame	5,242	28.6
1.0=the household did not participate (CAPI)	5,338	29.12
2.0=the household participated but informed consent NOT given	261	1.42
2.0=the household participated (CAPI started)	7,488	40.85

The variable A10X_PQ_PARTIMEN10A attempts to characterize the causes of nonparticipation (it represents the status of the last call attempted to contact a household that did not participate in the CATI). This variable only has a value for eligible nonparticipating households.

A10X_PQ_PARTIMEN10A	Frequency	Percentage
Other	88	3%
STOP REFUS 10,5 ans seulement (Refusal age 10½ only)	83	3%
STOP PARENT IMPOSSIBLE À INTERROGER (Parent could not be interviewed)	121	4%
ABANDON DU QUESTIONNAIRE (Left the questionnaire before completion)	145	4%
STOP REFUS DE NOUS PASSER LE PARENT (Refusal to give phone to parent)	164	5%
RDV (appointment not honored)	358	11%
STOP ABANDON ELFE (Refusal age 10½ and next waves)	412	13%
FAUX NUMÉRO (Wrong number)	512	16%
NRP/OCCUPÉ (NR/Busy)	1397	43%

Finally, the variables A010X_REFCOMP10A_CATI and A010X_REF2COMP10A_CATI characterize the two parents' participation and the completeness of the CATI for each one.

	A010X_REFCOMP10A_CATI	A010X_REF2COMP10A_CATI
0=Not applicable – household not participating	8,522	8,522
1=REF questionnaire complete	9,663	613
2=REF questionnaire incomplete	144	21
3=REF questionnaire not administered		222
4=Not applicable, no REF2 (variable <i>efvit</i> ne 4)		8,951

Weighting

The aim of the weighting presented here is to allow data from survey respondents to be used to infer characteristics of the target population (in the case of the ELFE cohort, the population of inference is that of infants born in 2011 in metropolitan France; at most a twin birth; excluding highly premature babies; with an adult mother)

The total size of the target population is set at 753,500 families (and 764,000 children). In the ELFE surveys, to correct for biases introduced by nonparticipation/nonresponse (avoiding the assumption that participants and nonparticipants have identical characteristics), rereighting is used.

Contrary to the preceding survey waves, only a cross-sectional weighting is proposed here, no longitudinal weighting (i.e. for participants in all survey waves).

This is because at this stage, there is a risk of the composition of the sample of children whose parent(s) participated in all waves no longer being representative of all children born in 2011. In addition, these longitudinal weights would be used only for analyses of variables covering all survey waves, which is unlikely. Thus, rather than calculating a set of weights that would be difficult or even impossible to use – with certain subpopulations so sparsely represented that extreme weights would result, while truncating these weights would risk strongly biasing the resulting analyses – the ELFE team has chosen not to provide these weights. It will of course be possible to produce specific weights upon request if needed (such as populations of children participating in a specific combination of waves – by example maternity, 3.5 years and 10.5 years only - and not all the waves between maternity and 10 years).

CATI (telephon survey)

In the case of the survey wave at age 10½, for all children, the first contact was a CATI survey of their parent.

We thus chose to begin the calculation of weights using only data from children with at least one parent who responded completely to a CATI questionnaire, a total of 9,663 children. The first set of weights made available for these children is the cross-sectional CATI weighting (A10E_PONDREF_CATI).

First of all, a logistic regression model was used to identify variables that contributed significantly to no-participation at CAT1 10 yrs (those without complete responses to the CATI at age 10½, out of the children included in the ELFE cohort in maternity units, a total of 18,329 children).

The following variables were chosen (*p*-value in red):

Effet	DDL	Khi-2	Pr > khi-2
		de Wald	
Region of maternity unit	4	26,814	<,0001
Primiparous mother	1	22,9195	<,0001
Marital status	1	4,3117	0,0379
Mother's age	3	111,7445	<,0001
Mother's level of education	3	535,0259	<,0001
Mother's immigrant status	1	128,167	<,0001
Birth preparation sessions	1	302,1658	<,0001
Father's activity status at birth	1	22,9876	<,0001
Mother cohabiting with partner at birth	1	52,5733	<,0001
Father's age	3	31,1643	<,0001
Alcohol consumption during pregnancy	1	30,0111	<,0001
Mother's activity status at the time of childbirth	1	96,6446	<,0001
Sex	1	0,8309	0,362

Only the child's sex was not significant. The included variables are thus:

- Regional group of residence, out of 5: Ile-de-France, Centre, Picardie; Northeast; Northwest; Southeast; Southwest
- Primiparous mother (yes/no)
- Marital status (parents married/unmarried at the time of birth)
- Mother's age (18-24/25-29/30-34/35+)
- Mother's level of education (no schooling, primary, lower secondary vocational [CAP], secondary vocational [BEP]/three last grades of secondary school/higher education)
- Mother's immigrant status (yes/no)
- Birth preparation sessions (yes/no)
- Father's activity status at the time of birth (in employment/other)
- Mother living with partner at birth (yes/no)
- Father's age (18-24/25-29/30-34/35+)
- Alcohol consumption during pregnancy (yes/no)
- Mother's activity status at the time of childbirth (in employment/other)

Twenty homogeneous participant groups (HPG) were constituted (with the probability continued participation in this survey wave ranging from 12.3% to 75.5% of the initial number of participants in the group).

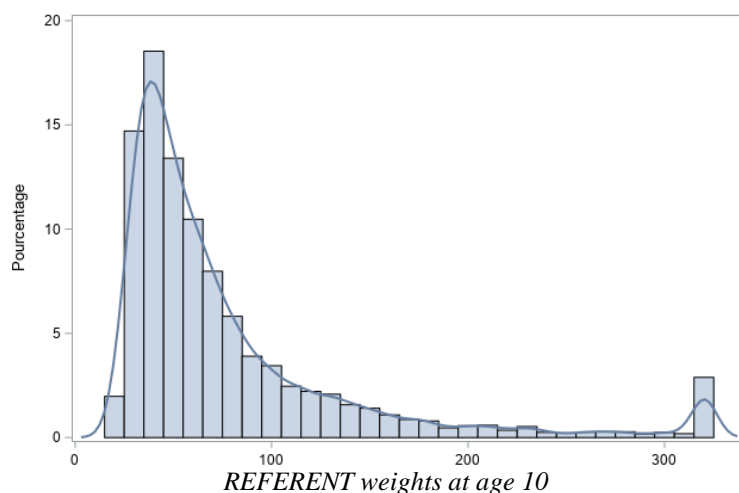
The initial weight was thus constructed by multiplying the weights from the selection (which takes into account only the size of the maternity unit in the 5-stratum classification and the season of birth) by the inverse of the probability of participation of the HPG to which the participating child belongs.

This initial weight was then calibrated using INSEE's CALMAR method, based on the 12 previously selected variables, with the addition of the child's sex and twin status (yes/no).

For these new variables, missing data (5% for father's age, below 2% for the rest) were imputed before the calibration. All of the totals for the ELFE population are drawn from vital records or the 2010 National Perinatal Survey.

Finally, extreme weights (those more than two standard deviations above the mean – a value of 300, a total of 3.5% of all weights – were truncated), and the set of weights was then adjusted to maintain a total of 764,000 children.

N	Mean	Max	Min	Intervalle	Sum	5%	10%	25%	50% median	75%	90%	95%	99%
9 663	79,06	320,52	19,48	301,04	764 000	27,97	31,39	39,22	56,20	89,90	155,23	231,10	320,52



CAPI (home survey)

In the case of the survey wave at age 10½, households that were reached for the telephone survey (even if they were not considered to have completed the questionnaire) were asked if they would be willing to have an interviewer visit them at home to ask the child additional questions. We thus calculated weights for children who responded to the CAPI and whose parents provided consent for a home visit (7,488 children). The weighting provided for these children is once again a cross-sectional one (A10E_PONDREF_CAPI).

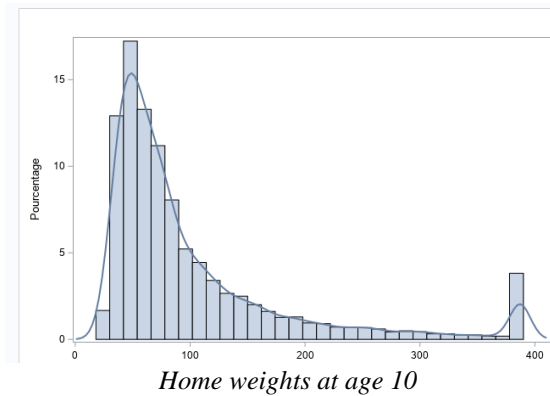
As for the CATI, a logistic regression model was used to identify variables that contributed significantly to the process of nonparticipation (non-completion of the CAPI at age 10½, among the 18,329 children included in the ELFE cohort in maternity units). The same significant variables were used as for the CATI:

- Regional group of maternity, out of 5: Ile-de-France, Centre, Picardie; Northeast; Northwest; Southeast; Southwest
- Primiparous mother (yes/no)
- Marital status (parents married/unmarried at the time of birth)
- Mother's age (18-24/25-29/30-34/35+)
- Mother's level of education (no schooling, primary, lower secondary vocational [CAP], secondary vocational [BEP]/three last grades of secondary school/higher education)
- Mother's immigrant status (yes/no)
- Birth preparation sessions (yes/no)
- Father's activity status at the time of birth (in employment/other)
- Mother living with partner at birth (yes/no)
- Father's age (18-24/25-29/30-34/35+)
- Alcohol consumption during pregnancy (yes/no)
- Mother's activity status at the time of childbirth (in employment/other)

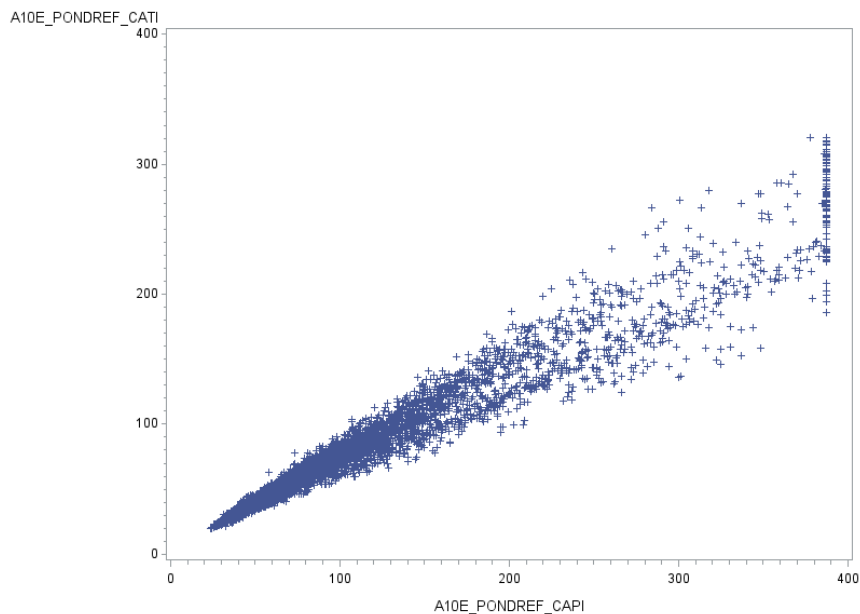
Twenty homogeneous participant groups were constituted, and the initial weights were then calibrated using INSEE’s CALMAR method, based on the 12 previously selected variables, with the addition of the child’s sex and twin status (yes/no).

Here again, extreme weights (those more than two standard deviations above the mean – a value of 350, 3.3% of all weights – were truncated). The weights were then adjusted to maintain a total of 764,000 children.

N	Mean	Max	Min	Intervalle	Sum	5%	10%	25%	50% median	75%	90%	95%	99%
7 488	102,03	387,35	23,39	363,95	764 000	34,25	38,74	48,93	70,67	117,48	213,98	318,16	387,35

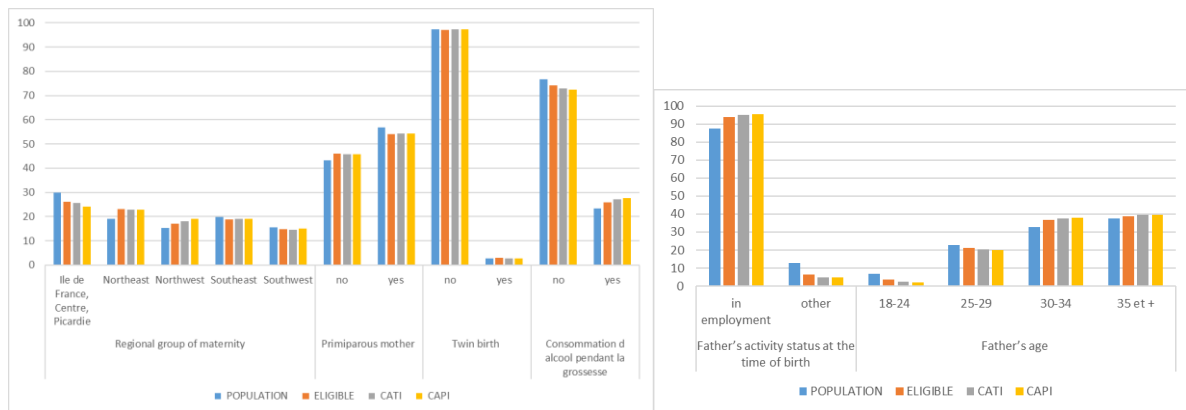


As the same variables were used to produce the two sets of variables, and as there are few major differences between the two samples of respondents, the weights in A10E_PONDREF_CAPI and A10E_PONDREF_CATI are highly correlated.

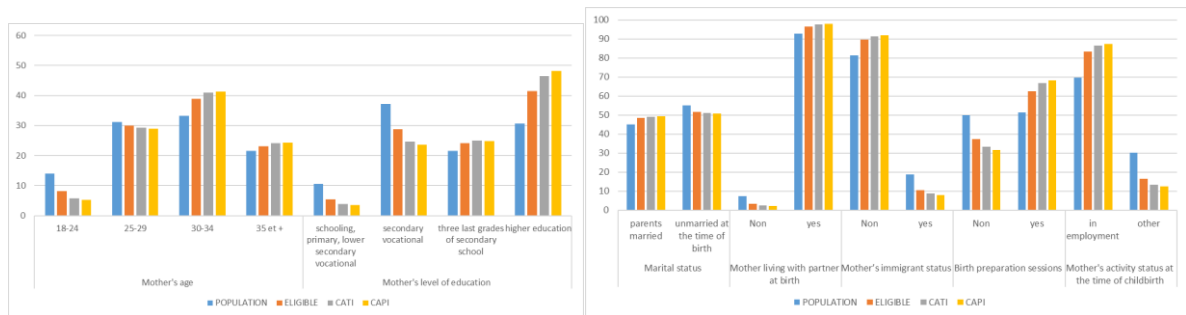


Who are the respondents to this survey wave?

To try and characterize the respondents to the CATI and CAPI surveys at age 10½, we can compare the distribution of all of the calibration variables across: the target population (measured as equal to the calibration margins); the population eligible for the survey (that is, those who were still followed at this age – to produce estimates on those who were lost to follow-up before this wave); and respondents at age 10½ (in order to estimate specific attrition in this survey wave).



Comparison of different populations – data on family and father



Comparison of different populations – data on mother

A few well-known phenomena can be seen in the graphs:

- There are particular populations with a significant proportion of nonparticipants from the outset: young parents, parents with low levels of education, immigrant parents, and unemployed parents. These populations left the survey over time (at the time of the recruitment or over time, as attrition accelerated).
- Finally, a certain population left the survey in the final phase at age 10½ (where the grey and yellow bars are shorter than the orange bar).

To attempt to estimate this recent attrition, we can concentrate on the differences between the eligible population and those who accepted to participate in the survey wave at age 10½. Few major differences are seen.

		ELIGIBLE	CATI			ELIGIBLE	CATI
Regional group of maternity	Ile de France, Centre, Picardie	26,1	25,6	Marital status	parents married	48,4	49,0
	Northeast	23,2	22,8		unmarried at the time of birth	51,7	51,0
	Northwest	17,1	18,0	Mother living with partner at birth	Non	3,4	2,5
	Southeast	18,8	19,1		yes	96,6	97,5
	Southwest	14,7	14,6		Mother's immigrant status	Non	89,5
Primiparous mother	no	46,0	45,7	yes		10,5	8,7
	yes	54,1	54,3	Mother's age	18-24	8,1	5,7
Twin birth	no	97,1	97,2		25-29	30,0	29,3
	yes	3,0	2,8		30-34	38,9	40,9
Consommation d'alcool pendant la grossesse	no	74,1	72,9		35 et +	23,1	24,1
	yes	25,9	27,1		Mother's level of education	schooling, primary, lower secondary vocational	5,5
Father's activity status at the time of birth	in employment	93,8	95,0	secondary vocational		28,8	24,7
	other	6,2	5,0	three last grades of secondary school		24,2	25,0
Father's age	18-24	3,6	2,5	higher education		41,5	46,4
	25-29	21,2	20,5				
	30-34	36,7	37,6				
	35 et +	38,6	39,4				

Note, finally that the proportions of the population represented by (the weights assigned to) the 9,160 children whose reference parent responded to the survey waves at ages 5½ et 10½ are quite strongly correlated. This ensures a certain continuity in analyses of the survey data: variability between measurements at the ages of 5 and 10 can be understood to result from variability in the phenomenon, and not from large variations in the weights between waves).

