

EESTI RAHVASTIKUSTATISTIKA  
POPULATION STATISTICS OF ESTONIA

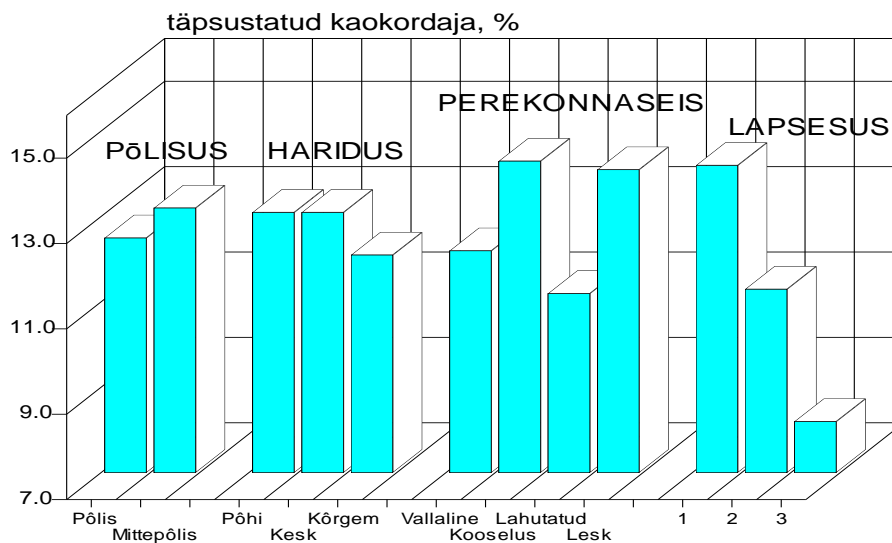
---

**EESTI PERE- JA SÜNDIMUSUURING**

Metodoloogiaülevaade

**ESTONIAN FAMILY AND FERTILITY  
SURVEY**

Methodological Report



Tallinn 1995



EESTI KÕRGKOOULIDEVAHELINE DEMOUURINGUTE KESKUS  
ESTONIAN INTERUNIVERSITY POPULATION RESEARCH CENTRE

# **EESTI PERE- JA SÜNDIMUSUURING**

Metodoloogiaülevaade

# **ESTONIAN FAMILY AND FERTILITY SURVEY**

Methodological Report

koostatud  
prepared by

Kalev Katus Allan Puur Luule Sakkeus Brian Silver

RU Seeria A No 39

Tallinn 1995

(C) Eesti Kõrgkoolidevaheline Demouuringute Keskus  
Estonian Interuniversity Population Research Centre

ISBN 9985-820-24-X



Eesti Kõrgkoolidevaheline Demouuringute Keskus  
Estonian Interuniversity Population Research Centre  
Postkast 3012, Tallinn EE0090, Eesti

Käesolev kogumik on autorikaitse objekt. Paljundusõiguse valdaja eelneva kirjaliku nõusolekuta on keelatud seda väljaannet või selle mistahes osa reprodutseerida, avaldada või jätta avaldamiseks infovõrgus, ümber kirjutada mistahes viisil või vahendiga elektrooniliselt, mehhaniliselt, fotokopeerimise, salvestamise või muul teel.

# SISUKORD

1. EPSU PÕHIJONED .....	5
1.1. Keskkond .....	5
1.2. Korraldus .....	7
1.3. Uuringuprogramm .....	8
1.4. Küsitluskogum .....	9
1.5. Rahastamine .....	10
2. KÜSITLUSINSTRUMENTAARIUM .....	11
2.1. Eluloosündmuste määratlused .....	11
2.1. Küsimustiku ülesehitus .....	11
2.2. Prooviküsitlused .....	13
3. VALIM .....	14
3.1. Üldkogum ja valimialus .....	14
3.2. Valimikava .....	14
3.3. Valimietapid .....	15
3.4. Küsitlustulemus .....	16
3.5. Kaopõhjused .....	17
3.6. Vastamiskordajad .....	18
3.7. Valimi esinduslikkus .....	18
3.8. Kooskõla loenduskirjetega .....	21
4. KÜSITLUS .....	22
4.1. Küsitlajaõpe .....	22
4.2. Küsitluskorraldus .....	23
4.3. Küsitlajate arv .....	24
4.4. Küsitluskuu .....	25
4.5. Küsitluskoht .....	26
4.6. Küsitluskestus .....	26
4.7. Küsitluskvaliteet: intervjuerijahinnangud .....	27
5. KODEERIMINE, SISESTUS JA ANDMEPUHASTUS .....	28
5.1. Küsimusjärgne vastuskadu .....	29
5.2. Küsitluskvaliteet vanemates põlvkondades .....	31
5.3. Andmefail .....	32
VIITED .....	34
LISAD	
A. EPSU instrumentaarium (eestikeelne versioon)	
B. EPSU instrumentaarium (venekeelne versioon)	
C. EPSU instrumentaarium (inglisekeelne versioon)	

# CONTENTS

1. EFFS PROJECT OUTLINE .....	36
1.1. Survey Environment .....	37
1.2. Management and Organisation .....	38
1.3. Programme .....	39
1.4. Survey Population .....	41
1.5. Funding .....	42
2. SURVEY INSTRUMENT .....	43
2.1. Definitions of Events .....	43
2.2. Questionnaire Design .....	43
2.3. Pretests .....	44
3. SAMPLING PROCEDURES .....	45
3.1. Target Population and Sample Frame .....	45
3.2. Sampling Plan .....	46
3.3. Stages of Sample Selection .....	47
3.4. Final Disposition of Cases .....	49
3.5. Reasons for not Completing Interviews .....	49
3.6. Response Rates .....	50
3.7. Representativeness of the Sample .....	51
3.8. Linkage with Census Records .....	55
4. DATA COLLECTION .....	55
4.1. Interviewer Training .....	56
4.2. Organisation of the Fieldwork .....	56
4.3. Number of Interviewers .....	58
4.4. Month of Interview .....	59
4.5. Location of Interview .....	59
4.6. Length of Interview .....	59
4.7. Quality of Interviews: Interviewer Assessments .....	61
5. CODING, DATA ENTRY AND EDITING .....	62
5.1. Item-Specific Non-Response .....	63
5.2. Quality of Interviews for Older Cohorts .....	65
5.3. Data File .....	66
REFERENCES .....	67
APPENDICES	
A. EFFS Survey Instrument (Estonian version)	
B. EFFS Survey Instrument (Russian version)	
C. EFFS Survey Instrument (English version)	

# SISSEJUHATUS

## 1. EPSU RIIGIUURINGUNA

Eesti Pere- ja Sündimusuuring (EPSU) on rahvuslik osaprojekt Euroopa Pere- ja Sündimusuuringu üldnime all tuntud rahvusvahelisest uuringusarjast, mida 1990 aastakümnel koordineerib ÜRO Euroopa Majanduskomisjoni Rahvastikuosakond (UN ECE PAU). EPSU ettevalmistamine algatati Eesti Demograafia Assotsiatsiooni ning Eesti Kõrgkoolidevahelise Demouuringute Keskuse (EKDK) poolt 1991 aastal. Kuivõrd EPSU on esimene omataoline uuring Eestis, tuli EPSU nõuetekohaseks läbiviimiseks lisaks tavapärastele uuringuülesannetele tegelda nõukogudejärgse andmekorralduse parandamisega üldisemalt. Vastavalt kujunesid kvaliteettulemuse saavutamiseks tarvilikud tegemised hoopis laiemaks kui ühe konkreetse küsitluse läbiviimine. Teisest vaatenurgast tagab rahvusvahelises koostöös osalemine EPSU kontseptuaalse ja metodoloogilise ühilduvuse rahvastikuteaduse kaasaegsete suundumustega ning avaldab soodsat mõju demograafia arengule Eestis. Samalaadiliste uuringute olemasolu enamikus Euroopa riikides kindlustab võimaluse määrata Eesti asend Euroopa ja eelkõige Baltoskandia rahvastikuarengu kontekstis.

### 1.1. Keskkond

Viimase poolsajandi vältel Eestis domineerinud totalitaarsel ühiskonnal oli teiste avaldumisvormide kõrval ka kindel spetsiifika andmekorralduse valdkonnas. Ehkki ühiskondlik olukord on vahepeal muutunud, toimib see omapärane andmekorralduse süsteem oma põhijoontes tänini [vt Anderson, Katus, Silver 1994]. Mõistmaks EPSU ettevalmistamise ja läbiviimise kulgu, aga samuti küsitlusuuringu korraldamisprobleeme post-sovjetlikus miljöös üldisemalt, on tarvilik alustada uuringukeskkonna käsitlusest. Rõhutamist väärivad selles osas neli põhilist tunnusjoont.

EPSU uuringukeskkonna esimeseks tunnusjooneks on selliste valimiprotseduuride olematus, mis on nõutavad esinduslike üleriigiliste küsitlusuuringute läbiviimiseks. Kuigi NLIidu korraldati perioodiliselt liiduvabariike hõlmavaid demograafilisi küsitlusi [Darsky 1986], koostati nende uuringute jaoks tarvilikud valimid eranditult Moskvas, NLIidu Statistika Keskvalitsuses. Ehkki ülalmainitud liiduliste uuringute valimimenetluste üksikasjalikku kirjeldust pole teadaolevalt publitseeritud, on ilmne, et nende tulemused pole olnud esinduslikud väiksemate liiduvabariikide, sealhulgas Eesti suhtes, isegi representatiivsuse korral üleliidulises ulatuses. Piiratud valimimahu ning mitmeastmelise klasterprotseduuri tõttu sattusid Eestist tavapäraselt valimisse vaid üksikud linnad ja maapiirkonnad. Tsentraliseeritud statistikakorralduse juures ei kujunenud vabariikide statistikainstitutsioonides vajadust omada küsitlusuuringute valimialust ega asjakohaseid oskusi. Arhiivimaterjalide kohaselt ei ole ka Eesti Statistikaamet kunagi omapoolselt tõstatanud küsimust nimetatud uuringute representatiivsusest vabariigi tasandil.

Teadlaskonnal oli nõukogulik andmekorralduses juurdepääs üksnes agregeeritud rahvastikuandmetele, valimialusena mõeldavate individuaalandmete mistahes kasutamine oli akadeemilistele institutsioonidele hoopis välistatud. Seega, omades isegi põhjalikke teoreetilisi

teadmisi valimiteooria valdkonnas, puudus teadlastel võimalus neid uuringute plaanimisel praktiliselt rakendada. Tulemusena osutus, rääkides Eestist, põhiliselt Tartu Ülikooli statistikute poolt arendatud valimilähenemine reaalsuskaugeks ning rakendamatuks kui valimialuseks sobiv andmestik sai kättesaadavaks. Taolises olukorras kujunes EPSU valimi koostamine standardprotseduuride rutiinse rakendamise asemel pigem innovaatiliseks protsessiks, mille käigus seati korda üleriigilisteks esindusuuringuteks tarvilik valimialus, viimistleti väljavõtumenetlus ja aadresside täpsustusprotseduur.

Teiseks EPSU uuringukeskkonna iseloomulikuks jooneks oli professionaalse üleriigilise küsitlajakõrgu puudumine. Nõukogude ajal oli Statistikaametil küll olemas oma intervjuerijate kaader perebühdzhetandmete kogumiseks, kuid ebapiisava professionaalse ettevalmistuse ning vähese paindlikkuse tõttu osutus see üleminekuaja algul tekkinud äriettevõtetele võrreldes vähesuutlikuks ja likvideeruks. Statistilistest küsitlustest lahus teostatud sotsioloogiliste uuringute läbiviimiseks mitmesuguste teadus- ja muude asutuste poolt kasutati reeglina ühekordselt palgatud küsitlajaid. Tegelikult pole professionaalse küsitlusvõrgu puudumises midagi ootamatut: kuivõrd Eestil ei olnud riigilise katvusega valimialust ega eeldusi esinduslikke uuringuid teostada, siis ei läinud tarvis ka vastavat võrku.

Püsiva küsitlajakõrgu olematuse tõttu puudus potentsiaalsetel intervjuerijatel, kes olid küsitlustööga varem kokku puutunud ning keda saanuks EPSU puhul põhimõtteliselt kasutada, süstemaatiline väljaõpe. EPSU ettevalmistamise käigus ilmnis pahatihti küsitlustöö põhindoete vähenemine tundmine, kõnelemata delikaatseid elutahke puudutava sündmusloolise andmestiku kogumiseks tarvilikust kvalifikatsioonist. Usaldusväärse töö kvaliteediga võrgu puudumine muutis intervjuerijate väljaõppe ning nende töö koordineerimise üheks kõige kesksemaks ülesandeks EPSU ettevalmistamisel. Andmekvaliteedi tagamiseks tuli uurimiserühmal sellega tegeleda töömahukusele vaatamata välitööde kõigil etappidel.

Kolmas keskkondlik probleem, millega EPSU ettevalmistamine ja teostamine kokku puutus, oli projektijärgse rahastamispraktika puudumine riigieelarves. Osaliselt peegeldab taoline olukord rahvastiku-, iseäranis aga küsitluste teel hangitud rahvastikuandmete üldist mitteasjakohasust soliidsete riigiasutuste töös. Iseloomulik selles osas on Statistikaameti poolt veel hiljaaegu väljendatud arusaam, mille kohaselt igasugune individuaalandmetega tegelemine asub väljaspool riigistatistika huviväljaala. Tausta arvestades mõjus EPSU-taolise suurküsitluse eelarvesse lülitamine piisavalt uudsena. Aastatsükliga toimivas eelarvemenetluses puudusid protseduurid, mille abil oleks saanud esitada taotlust ning langetada otsust kogu uuringu osas tervikuna. Projekti rahastamise seisukohalt ei mõjunud soodsalt ka rahareformile eelnenud järsk rublainflatsioon ning üleminekumajandusele iseloomulikud finantspiirangud. Antud oludes nõudis asjaosalistelt projekti rahastamine suurt selgitus- ja *lobby*-tööd ametlikes ringkondades. Nii näiteks oli EPSU korduvalt arutusel Riigikogu Eelarve- ja Majanduskomisjonis, kolmel korral otsustati uuringu kui omaette hääletusobjekti üle parlamendi täisistungil. Lõpptulemusena eraldati uuringu tarbeks vahendeid ühe kolmandiku ulatuses koguvajadusest, mis muutis hädatarvilikuks rahvusvahelise rahataotlemistegevuse.

Neljas EPSU uuringukeskkonna tähtsam tunnusjoon on seotud rahvastikuandmete võrreldavuse ja kvaliteediga. Erinevalt mitmest teisest endise NLiidu osast, viidi Eestis nõukogude ajal läbi arvestatav hulk sotsioloogilisi küsitlusi. Ehkki enamjaolt spetsiifilistele kogumitele ja/või teatud eriprobleemidele keskenduvad, oli nende hulgas ka arvestatava ulatusega projekte (näiteks, Sotsioloogia Instituudi poolt läbiviidud noorsookohortide longituuduuring [Titma 1993] ja Tallinna Linnauurimisinstituudi poolt perioodiliselt teostatud Tallinna elanikkonna kompleksküsitlused [Pavelson, Mihhailov 1982]). Tüüpiliselt kasutasid tollased uuringud



valimialusena ettevõtte/asutuse töötajaskonna või siis valijate nimekirju, millele on iseloomulik respondente puudutava taustinformatsiooni nappus ning töötajate nimekirjade puhul ka ebaadekvaatne kaetus. Kuivõrd vähesed tollastest uuringutest omavad dokumenteeritud metodoloogiaülevaateid, siis ei saa neid reeglina võrdlevanalüüsil kasutada, ehkki andmed ise on Eesti Raadio arvutuskeskuses säilitatud.

Võrreldavusaspektist on need andmekogumid omaetteiseisvad, mida pole võimalik ühendada ei omavahel ega ka muude andmepankadega. Taolise perspektiivi vältimiseks on EPSU puhul algusest peale püütud mõelda kontseptuaalsele haakuvusele teiste kesksimate Eesti andmepankadega. Ehkki Statistikaamet suhtus taolisesse püüdlustesse pigem kahtlusega, on ühilduvus tänaseks õnnestunud kindlustada 1989 aasta rahvaloendusega, Eesti Tööjõu-uuringuga ning sündmusstatistika andmestikuga sünni- ja surmajuhtude osas.

Järelnõukoguliku uuringukeskkonna kirjeldatud tunnused tõid kaasa hulga igapäevaseid lisäülesandeid, mille lahendamine normaaloludes ühe konkreetse uuringu raamistikku ei kuulu ning millela uuringu tulemused oleksid ehk kiiremini kasutajani jõudnud.

## **1.2. Korraldus**

Keskkonnaolusid arvestades kujunes EPSU eduka teostamise eeltingimuseks kõigi asjasthuvitatud institutsioonide ning üksikuurijate jõupingutuste koondamine. Eesti Demograafia Assotsiatsioonis toimunud arutluste tulemusena otsustati projekti ettevalmistamiseks ja läbiviimiseks moodustada Teadusnõukogu ning anda vahetud tööfunktsioonid Eesti Kõrgkoolidevahelisele Demouuringute Keskusele.

Ametlikult loodi teadureid ning riigiasutuste asjatundjaid ühendav Teadusnõukogu mais 1992. Kutsed selle töös osaleda saadeti kõigile asjaomastele institutsioonidele, kelle esindajad moodustavadki Teadusnõukogu liikmeskonna: Hans Hansen (Tallinna Pedagoogikaülikool), Kalev Katus (esimees, EKDK), Aili Kelam (Sotsioloogia Instituut), Toivo Kitvel (Tallinna Pedagoogikaülikool), Jüri Kruusvall (Tallinna Pedagoogikaülikool), Jüri Kõre (Tartu Ülikool), Dmitri Mihhailov (Vene Kultuuriselts), Katrin Paadam (Tallinna Tehnikaülikool), Allan Puur (EKDK), Asta Põldma (sekretär, Eesti Demograafia Assotsiatsioon), Leena Rõbakova (sekretär kuni 1993, Eesti Demograafia Assotsiatsioon), Luule Sakkeus (EKDK), Helbe Sinimäe (Tartu Ülikool), Mati Sundja (Riigi Statistikaamet), Lee Tammemäe (Perinatoloogia Selts), Lembit Tepp (Riigi Statistikaamet), Andres Vikat (EKDK), Rein Võormann (Sotsioloogia Instituut). Hiljem asendas Hill Kulu (Tartu Ülikool) Jüri Kõre. Väliseksperdist liikmetena on Teadusnõukogu tegevuses osalenud Barbara A. Anderson (Michigani Ülikool), Brian D. Silver (Michigani Riiklik Ülikool) ja Jan M. Hoem (Stockholmi Ülikool).

Alates ettevalmistusetapist on Teadusnõukogu kandnud vastutust kogu EPSU uuringuterviku eest, mis jätkub ka käesoleval ajal analüüsi koordineerimise vormis. Sisuliste küsimuste kõrval on Teadusnõukogu pädevuses otsustada ka uuringu läbiviimiseks eraldatud rahaliste vahendite kasutamise üle. Keskastme administratiivbürokratiast rippumatuna on Teadusnõukogu toiminud iseseisva üksusena Isikuandmenõukogu pädevusalas.

EPSU vahetuid korraldusfunktsioone on selle kõigil etappidel täitnud Kõrgkoolidevaheline Demouuringute Keskus. Aastatel 1992-1993 töötas EKDK Teadusnõukogu juhendamisel välja küsitlusinstrumentariumi, valmistas ette valimialuse ning moodustas valimi. Läbiviidud kolme

prooviküsitluse põhjal valis Teadusnõukogu välitööde läbiviijaks Saar Polli küsitlajakavaga. 1994 aasta algul viidi EKDK poolt läbi intervjuerijate ankeedikeskne väljaõpe, intervjuud toimusid sama aasta jaanuarist novembrini. 1995 aastal viidi lõpule andmesistestus ning andmepuhastus. Uuringu esmaspublikatsioonidena ilmuvad 1995 aastal *Standardtabelid* ning *Metodoloogiatähelepanu*. Analüüsijärku jõuab uuring 1996 aastal.

Ametlike struktuuride väljaarenemise tõttu on EPSU Teadusnõukogul tulnud täita mitmeid niisuguseid funktsioone, mis tavaoludes ühe uuringu töörühma kohustustesse ei kuulu. Tagantjärele tuleb iseäranis välitööde ja andmesistestuse juhtimist töömahukusele vaatamata lugeda uuringu eesmärkide saavutamise ning andmekvaliteedi tagamise seisukohalt mõõdapäasmatuks. EPSU käigus omandatud kogemusi on hiljem arutatud Isikuandmenõukogus ning osalt juba kasutatud järgnevate statistiliste tööde teostamisel, näidetena võib mainida Tööjõu-uuringut ning riigi küsitlajakavaga ülesehitamist maakondlike statistikabüroode baasil.

### **1.3. Uuringuprogramm**

Metodoloogiliselt rajaneb EPSU sündmusloolisel lähenemisviisil, mis keskendub inimese tähtsamate elusündmuste ajastusele ning järjestusele. Kaasaegses rahvastikuteaduses on sündmuslooline lähenemine vahendiks, mille abil seostatakse individuaaltasandi inimkäitumine ühelt poolt, ning rahvastikuprotsesside makromuutused teiselt poolt [Tuma, Hannan 1984; Blossfeld et al 1989 etc]. Teave elusündmuste ning nende toimumisaja kohta on kogutud tagasivaateliselt küsitlitava sünnist kuni intervjuuhetkeni. Põhiosas on uuringuprogramm kooskõlas UN ECE PAU tuumküsimustikuga, sisaldades küsimusmooduleid vanematekodu, kooselude ja abielude, raseduste ja sündide, elukohavahetuste, haridus- ning tööteede kohta [UN ECE 1992b]. Võimaldamaks sissevaadet sõjajärgsesse Eesti rahvastiku arengusse ning loomaks tarvilikku alusandmestikku rahvastiku- ja sotsiaalpoliitika tarvis, on uuringu programm mitmes probleemivaldkonnas tuumküsimustikuga võrreldes põhjalikum.

Üheks laiendatud käsitlusega valdkonnaks on rasedus- ja sünnilugu. Tarvidus sügavamale teabe järele on siinkohal tingitud Eesti väga kõrge rasedus- ja abortiivsustasemest: viimaste aastakümnete jooksul on abortiivsus püsivalt ületanud sündimust [Anderson, Katus, Puur, Silver 1993]. EPSU peaks tooma selgust nende sotsiaalsete ja individuaalsete tegurite osas, mis kujundavad valiku sünnituse ja aborti vahel. Et enamus esmasrasedusi saab alguse väljaspool abielu, omandab rasedusväljundi seisukohalt suure tähtsuse esmaskooselu. Ajastusest kõneldes kaasnes mitme kümnendi vältel menstruatsiooni, seksuaalelu ning esmaskooselu noorenemisega samasuunaline sünniajastuse trend. Märkides järjekordset pöördepunkti rahvastikuarengus, niisugune ühesuunalisus 1970 aastate lõpul katkes: sündimustrend omandas varasemaga võrreldes hoopis suurema iseseisvuse.

Eestis, sarnaselt teiste modernse taastetüübiga rahvastikega on taastetase esmajoones sõltuv kolmandate sündide osakaalust põlvkonnas. Sündmusstatistika andmetel on Eesti põlisrahvastikul kolmanda lapse sünnitanud naiste osakaal olnud põlvkondade asendumiseks piisav ajavahemikul 1968-1990; varem ega ka hiljem pole demograafilise ülemineku järgsel perioodil taolist soodsat olukorda esinenud. Arvestades taasteprobleemi keskset tähendust, peaks järjestus-spetsiifiline lähenemine abielu-, seksuaal-, abortiiv- ning sündimuskäitumisele välja selgitama need eluteed, mille läbi kolmelapselise pereni Eestis jõutakse.

Teiseks uuringus sügavamalt käsitletud pälvinud valdkonnaks oli rände- ja eluasemelugu. Teadaolevalt on Eesti kogu sõjajärgse demoarengu üheks põhijooneks olnud välispäritolu rahvastikuosa järsk suurenemine: viimase rahvaloenduse andmeil ületas välissündinute osakaal rahvastikus veerandi, koos sisserännanute teise põlvkonnaga aga ligines koguni kahe viiendikuni kogurahvastikus [Sakkeus 1991; 1994]. Võrrelduna põlisrahvastikuga, iseloomustavad immigrante noorem vanuskoostis ning põlisrahvastikust erinevad demograafilised ja sotsiaalsed käitumistavad. Sisserändes eristuvad rändetagamaa laienemisest tingituna kaks suuremat migratsioonilainet, mis ühtlasi on Eesti välissündinute terava seesmise diferentsiatsiooni aluseks.

Kõrvuti rändelise rahvastikuosa järsu suurenemisega on Eestile viimase poolsajandi vältel olnud iseloomulik ka intensiivne sisemigratsioon, mille üheks osaks on 1940-1950 aastate sundränne. Siseliikumise tulemusena on toimunud märgatavad nihked rahvastiku paiknemises ning kujunenud iseloomuliku rahvastikupildiga sisse- ja väljarändepiirkonnad. Nende tekkemehhanismide jälgimisele lisaks pakub EPSU andmestik võimalusi lühi- ja pikaajalise, samuti lähi- ja kaugrände rahvusvaheliselt võrreldavate definitsioonide rakendamiseks. Elukohavahetuste ning peresündmuste ajastuse kõrvutamise aitab eeldatavalt selgust tuua rände ning perekarjääri omavahelistesse seostesse.

Eelnimetatud valdkondadele lisaks on EPSU programm mõnevõrra laiendatud ka vanematekodu ning leibkonna ainelise kindlustatuse suunal. Teisalt osutusid osa hoiakuid ning arvamusi puudutavaid UN ECE PAU tuumküsimusi Eesti üleminekuolude kiire muutlikkuse tõttu eksitavaks ning tuli kõrvale jätta. Mõned väärtusorientatsioonid puudutavad küsimused olid aga Eesti konteksti arvestades rakendamatud. Tervikuna on EPSU küsimustik teiste riikidega hästi võrreldav, 1994 aastal UN ECE PAU poolt teostatatud analüüsi põhjal kuulus sellele Euroopa Pere- ja Sündimusuuringu ühilduvusskaalas kolmas koht.

#### **1.4. Küsitluskogum**

Põhijoontes järgib EPSU valimikava UN ECE PAU poolt osalejamaadele antud soovitusi. Uuringu üldkogumi moodustasid aastatel 1924-73 sündinud naised, kes 1994 aasta algusel olid jõudnud vanusvahemikku 20-69 eluaastat. Saavutamaks valimi võimalikult täielikku vastavust üldkogumile nii ühiskonnakihtide kui regionaalüksuste lõikes, kasutati valimi koostamisel üheastmelist lihtjuhuvalikut ilma klasterdamise ja geograafilise stratifikatsioonita. Nii sisuliselt kui finantskaalutlusi arvestades koostati esialgne valimikava kahevariandilisena. Miinimumvariandi kohaselt tulnuks intervjuuerida 5000 naist vanuses 20-69 aastat, maksimumvariandi puhul aga küsitleda 7500 naist ja 2500 meest; lisandunud 2500 naist pidanuks kuuluma sünnitusikka vanuses 20-49 aastat. Rahastamispiirangute tõttu leidis tegelikkuses realiseerimist valimikava miinimumvariant.

Esinduslikkuse seisukohalt langes küsitletute koostis küllalt täpselt üldkogumi struktuuriga kokku. On tähelepanuväärne, et valimisse kaasatud vanemate põlvkondade puhul polnud vastamiskordaja keskmisest madalam. Valimi ja üldkogumi koostise kokkulangevuse tõttu ei nõua EPSU andmestik tulemuste representatiivsuse tõstmiseks täiendavat kaalumist.

Valimialuse osas toetub EPSU 1989 aasta rahvaloenduse individuaalandmestikule, mis on tänaseni ainsaks allikaks, mis tagab üldkogumi kaetuse ning iga valimisse sattunud indiviidi piisava taustinformatsiooniga kindlustatuse. Ettevalmistus valimi koostamiseks algas 1991

aastal, mil Eesti Kõrgkoolidevaheline Demouuringute Keskus koostöös Eesti ja Venemaa Statistikaametite ning Michigani Ülikooli teadlastega kandis loendusandmestiku personaalarvuti jaoks sobivale kujule. EKDK poolt andmestik korrastati ja dokumenteeriti, mis oli eelduseks selle valimialusena kasutamisel [Katus, Puur 1993].

Ehkki EPSU valimikava vastas UN ECE PAU soovitudele, erines üldkogumi kaetus siiski ühes suhtes reeglipärasest. Nimelt oli küsitlevate vanuse ülempiiriks üldkasutatava 50 asemel 69 eluaastat. Vanuspiiri edasinihutamise tingisid kaks asjaolu. Esiteks, puuduvad Eestis varasemad sündmuslõlised uuringud ning vastavat teavet pole võimalik saada ka teistest allikatest. Seetõttu osutus vanemate põlvkondade uuringusse lülitamine ainsaks võimaluseks, kuidas kujunenud teabelünka täita. Teiseks tähtsamaks motiiviks vanemate naiste küsitlemisel oli informatsiooni hankimine Eesti sõjajärgse rahvastikuarengu olulisemate murdepunktide, eriti sündimustrendi osas.

Erinevalt enamikust Euroopa riikidest puudub Eesti sündimustrendis sõjajärgsele perioodile iseloomulik tõusulaine; sündimuse suurenemine leidis Eestis aset hoopis 1960 aastate lõpul, mil sündimustrend mujal Euroopas oli languseks pöördunud [Katus 1991; 1994]. Omamata teavet vanemate põlvkondade kohta, oleksid need Eesti rahvastikuarengu omapära avavad trendimuutused jäänud paratamatult käsitlematuks. Lisaks sündimusanalüüsile eeldab pikemaajalist informatsiooni ka teiste rahvastikuprotsesside käsitlemine. Arvestades välispäritolu rahvastiku mõju Eesti senisele ja tulevasele arengule, on teave kõigi immigratsioonikohortide kohta oluliseks abiks ühiskonnas aset leidvate integratsiooni- ja adaptatsiooniprotsesside tunnetamisel. On heameel tõdeda, et tänu küsitlusel rakendatud abinõudele ei toimunud vanemate põlvkondade valimisse lülitamine andmekvaliteedi halvendamise hinnaga.

## **1.5. Rahastamine**

EPSU rahastamine on toimunud mitmetest allikatest. Suurem osa uuringu ettevalmistamise ning välitöödega seotud kuludest on kaetud Eesti riigieelarvest. Ehkki eraldised 1993 aasta eelarvest katsid vaid kolmandiku projektile vajaminevast summast, on järgnevatel aastatel jätkunud rahastamine riigi osalust oluliselt suurendanud. Eelarvevahendite eraldamine uuringule toimus Statistikaameti kaudu, vahendite kasutamist jälgis Isikuandmenõukogu.

EPSU algetapil oli uuringu käekäigu seisukohalt otsustav tähtsus Kõrgkoolidevahelise Demouuringute Keskuse panusel. EKDK teadurid kindlustasid valimialuse, tagasid töökeskkonna, ennekõike aga eraldasid oma aja uuringu kõikide seniste tööetappide vältel. Mitmel korral võttis EKDK endale eelfinatseerimisrisiki.

EPSU väliseksperdid, Barbara Anderson ja Brian Silver, toetasid oma grantisummadega uuringu küsimustikule erimoodulite lisamist või süvendamist. Samad teadlased finantseerisid ka 1992 aastal läbiviidud abordialast prooviküsitlust (NICHD Grant Nos. HD-19915 ja P30 HD-10003). Väliseksperptide Andersoni, Hoemi ja Silveri hinnatavaimaks panuseks on Teadusnõukogu töökoosolekutel osalemisele pühendatud aeg ja energia, suur osa Eestiskäikude kuludest jäi samuti ekspertide endi kanda.

ÜRO Rahvastikufond (UNFPA) toetas uuringut välitööde lõpuleviimise, andmesisestuse ning -puhastuse osas, samuti *Standardtabelite* ettevalmistust ning *Metodoloogiaülevaate* koostamist

(Grant No EST/95/P01). UNFPA toetus oli äärmiselt oluline välitööde lõpuks kuhjunud rahastamisprobleemide lahendamisel. EPSU teadusanalüüsi alustamine, sealhulgas standardtabulatsioonide ning metodoloogiaülevaate analüütilised tööd on teostatud Eesti Teadusfondi toetusel (Grant No 1485).

Euroopa Pere- ja Sündimusuuringu rahvusliku osaprojektina on EPSU koordineerinud ÜRO Euroopa Majanduskomisjoni Rahvastikuosakonna poolt, metodoloogiliselt järgib uuring UN ECE soovitusi [1992a; 1992b; 1993]. Lisaks koordinatsioonile on UN ECE abistanud projekti nõuannete, oskusteabe vahendamise ning väljaõppe võimaldamisega, sealpoolne koordinatsioon jätkub eeldatavalt ka uuringu analüüsi etapil.

## **2. KÜSITLUSINSTRUMENTAARIUM**

### **2.1. Elusündmuste määratlused**

Eesti Pere- ja Sündimusuuring on eesmärgistatud koguma andmeid inimese elu kõigi tähtsamate sündmuste ja asjaolude, nimelt vanematekodu, sotsialiseerumiskeskonna, haridustee, sotsiaalse iseseisvumise, perekonnaloole, seksuaal- abortiiv-, kontratseptiiv- ja sündimuskäitumise, rändete, töö ja eluaseme kohta. Elusündmuste määratlused tuginevad rahvusvaheliselt kasutatavatele definitsioonidele. Asjaomaste andmedefinitsioonide kohandamine Euroopa Pere- ja Sündimusuuringule toimus osalejamaade koostöös, selle töö tulemus on kajastatud UN ECE PAU tuumküsimustikus [UNECE 1992b]. Andmete esitusmääratlused on toodud koordineeriva institutsiooni teises publikatsioonis [UNECE 1993]. On tähelepanuväärne, et kasutatud määratlused on samuti ühildatavad rahvusvaheliste loendus- ja sündmusstatistika definitsioonidega.

Kui andmedefinitsioonid olid kokku lepitud ühistöö tulemusena, siis konkreetse rakendusjuhendi koostamine jäi muidugi iga riigi iseseisvaks tööks. Sõltuvalt oludest võis ja pidigi rõhuasetus olema erinev, tagamaks lõpptulemusena võrreldava info küllaltki erineva rahvastikuarengulise ja sotsiaal-majandusliku olukorraga riikide kohta. Kuivõrd eestikeskne andmemetodoloogiline on alles algamas, kujunes Eesti olusid arvestava andmejuhendi koostamine vastutusrikkaks ja paljuski innovaatiliseks tööks. Küsitluse läbiviimine paralleelselt eesti ja vene keeles tingis ka andmejuhendi koostamise kahes keeles. Märkimisväärsed erisused rahvastikuarengus leidsid väljenduse juhendi eri versioonide rõhuasetustes. Käesoleva ülevaate koostajad peavad andmejuhendi juures tehtud tööd sedavõrd oluliseks, et see täielikul kujul lisas esitada.

### **2.2. Küsimustiku ülesehitus**

Oma ülesehituses järgib EPSU küsimustik Euroopa Pere- ja Sündimusuuringu tuumküsimustikku [UNECE 1992b]. Nagu eelpool mainitud, on Eesti olusid arvestavalt laiendatud uuringuprogrammi eelkõige rasedus- ja sünniloo ning rände- ja eluasemeloo, aga ka vanematekodu ning leibkonna ainelise kindlustatuse osas. Kokku koosneb küsimustik 13 osast:

1. Leibkond
2. Kooselud
3. Praegune elukaaslane
4. Lapsed
5. Teised rasedused
6. Sündimuse reguleerimine
7. Sündimushoiakud
8. Muud väärtushinnangud
9. Vanematekodu
10. Rände- ja eluasemelugu
11. Leibkonnalugu
12. Haridus- ja töölugu
13. Praegused elamistingimused

UN ECE PAU tuumküsimustikus algab eluloosündmuste fikseerimine kalendriaastast, mil küsitlaval täitus 15 eluaastat. EPSU ankeedis otsustati sündmuste fikseerimist alustada 14 eluaastast, kuivõrd uuringuga hõlmatud eakamatel põlvkondadel (mäletatavasti lisati küsitlusse naised vanuses 50 kuni 70 eluaastani) lõppes hariduskohustus 14 eluaastaks.

Küsitlustehniliselt on EPSU ankeet küllalt nõudlik, sisaldades hulgaliselt üleminekuid ning alternatiivsovnastusi, mis nõuavad intervjuerijalt paindlikku reageerimist ning vastaja elulookontekstis orienteerumist. Sündmusloolise andmestiku kogumine toimus vastavasisuliste tabelite abil, mille igasse veergu märgiti ühe, antud tüüpi elusündmuse kohta käiv teave. Küsitlaja täitis elulootabelid veerukaupa, fikseerides kõigepealt ühe, seejärel teise, siis kolmanda jne sündmuse. Kui sündmuste arv osutus suuremaks kui trükitud veergude arv, siis kasutas küsitlaja selleks otstarbeks ettenähtud lisalehe.

Tagamaks küsimustiku eri osades kogutud informatsiooni kooskõla, täitis intervjuerija konkreetsete küsimustega paralleelselt ankeedi lõpus olevat koondtabelit. Tabelis vastas igale kalendriaastale üks rida, iga sündmusliigi jaoks oli ette nähtud eraldi veerg. Sündmuste toimumisel tegi intervjuerija koondtabeli vastavate veergude ja ridade ristumiskohale tarvilikud märgid, mille abil pärast küsitluse lõppu oli lihtne saada kiire ülevaade eri eluvaldkondade omavahelisest haakuvusest. Tabeli peaesmärk oli anda küsitlajale võimalus kõrvaldada ankeedis olevad vastuolud enne intervjueritava juurest lahkumist.

Küsimustiku juurde kuulub mahukas intervjuerijajuhend, milles antakse juhised nii ankeedis käsitletavate protsesside määratluste kohta kui ka küsimustejärgsed selgitused ankeediga tegutsemiseks. Lisaks juhendile sisaldab EPSU instrumentarium veel protseduurijuhendi, vastusvariantidega kaardivihiku, ankeedi juurde kuuluvad lisalehed, küsitlusprotokoll, kontaktkirja ja vastuskaardid. Tähtsamad instrumentariumi osised on ära toodud käesoleva metodoloogiaülevaate lisas.

EPSU sihtrahvastiku rahvus- ja keelekoostise tõttu tuli kogu küsitlusinstrumentarium ette valmistada identsena kahes keeles. Instrumentariumi originaal valmis eesti keeles, misjärel see kohandati vene keelde. Mitmel juhul aitas venekeelse instrumentariumi, eriti ankeedi koostamine parandada ja täpsustada ka eestikeelset varianti. Ankeet ja küsitlajajuhend on tõlgitud ka inglise keelde.

### 2.3. Prooviküsitlused

EPSU küsimustiku, küsitlusprotseduuri ning potentsiaalsete küsitlusejavõrkude hindamiseks viidi põhiküsitluse eel läbi kolm prooviuuringut. Esimene prooviküsitlus toimus 1993 aasta veebruaris-märtsis Tallinnas ning Viljandi linnas ja maakonnas. Tallinnas viis küsitluse läbi küsitluskeskus *Saar Poll* (30 küsitletut) ning Viljandis kohalik Statistikabüroo (18 küsitletut linnas ja 20 maapiirkonnas). Esimese prooviküsitluse eesmärk oli testida küsimustiku eestikeelset varianti ning selgitada erinevusi linna- ja maaelanikest küsitletavate vastamises. Vanemate naiste tegelikust suurem esindatus prooviküsitluse valimis oli suunatud meenutusraskuste hindamiseks.

Eelkõige oli esimene prooviküsitlus aluseks küsimustiku täiustamisele. Ühtlasi näitas prooviküsitlus ilmekalt, et sündmusloolise andmestiku kogumine eeldab küsitlejate põhjalikku ettevalmistust ning head ankeeditundmist. Saadud kogemustele tuginedes koostati uuringu küsimustiku uus versioon, mida viimistleti koos väliskonsultantidega EPSU ekspertseminaril Tallinnas 1993 aasta mais.

Teine prooviküsitlus viidi läbi 1993 aasta juulis-augustis. Selle suundeesmärkideks oli testida 1989 aasta rahvaloendusandmete kasutusvõimalust valimialusena ning uuringu küsimustiku venekeelset varianti. Küsitlus viidi läbi Narvas, kus küsitlusperiood sattus ettekatsetamatult poliitiliselt pingelisele ajale (käimas oli referendum selle piirkonna Eestist eraldumise küsimuses). Prooviküsitlus andis kinnitust, et küsitletavate rahvaloendusaegseid aadresse saab Keskaadressbüroo abiga ajakohastada ning loendusandmestik on valimialusena igati kõlbulik. Küsitlus näitas samuti, et venekeelse elanikkonna intervjuerimine ei kujuta endast täiendavat probleemi: ükski 62-st valimisse sattunud ei keeldunud vastamast ning ainult kolme asukohta ei suudetud aadressi muutumise tõttu välja selgitada. Teise prooviküsitluse käigus katsetati ka Filosoofia, Sotsioloogia ja Õiguse Instituudi sotsioloogia osakonnal baseeruvat küsitlusvõrku.

Lisaks kahele eelnimetatule viidi Tallinnas EPSU ettevalmistamistöe ühe osana 1992 aastal läbi spetsiaalne prooviuuring abortiiv- ja kontratseptiivkäitumise küsitlusvastuste usaldusvääruse selgitamiseks. Uuring näitas, et Eestis suhtutakse reproduktiivvaldkonda puudutavatesse uuringuküsimustesse mõistvalt ning neile vastamine ei valmista enamikule respondentidest raskusi. Rohkem kui 80 protsenti küsitletutest, kellel oli uuringule eelnenud aasta jooksul olnud abort, tunnistas seda ka küsitlejale; USAs näiteks saadakse sedalaadi küsimustele vastus vaid 30-50 protsendi respondentide käest [Anderson, Katus, Puur, Silver 1993; 1994].

Kõigi kolme prooviküsitluse tulemusi arvestavalt valmis 1993 aasta oktoobris EPSU küsimustiku lõplik versioon. Teostatud võrdlusele tuginedes otsustas Teadusnõukogu välitööde läbiviimiseks kasutada *Saar Polli* küsitlusejavõrku.

### 3. VALIM

#### 3.1. Üldkogum ja valimialus

EPSU üldkogumi moodustasid naispõlvkonnad 1924-1973 ehk eluaastates 20-69 (1994) olevad naised. Valimisse sattumise võimalus oli võrdne kõigil nimetatud vahemikku mahtuvatel Eesti alalistel elanikel. Üldkogumi määratlemise järel tuli leida sobiv valimialus, mille põhjalt väljavõtt teostada. Nagu eelpool öeldud, sai selleks 1989 aasta rahvaloenduse individuaalandmestik. Ehkki loenduse kriitilise momendi näol mõnevõrra aegunud hetkeseisu peegeldav, oli loendusandmestik üldkogumit kõige terviklikumalt kattev andmebaas, mis võimaldas iga valimikirje varustada 55 erinevat tunnust hõlmava taustinformatsiooniga [Katus, Puur 1993]. Allpool on seda taustteavet rakendatud kaoanalüüsil ning tulemuste representatiivsuse ja kooskõla hindamisel. Muuhulgas välistas mitmekülgse küsitluseelse informatsiooni olemasolu ka fiktiivsete intervjuude võimaluse.

Loendusest möödunud ajavahemik tõi endaga kaasa teatud ala- ja ülekaetuse olemasolu valimialuse suhtes. Ühelt poolt sisaldas loendusandmestik isikuid, kes olid pärast 1989 aasta jaanuari Eestist lahkunud või surnud. Teisalt ei sisaldanud valimialus neid, kes olid loendusele järgnenud aja jooksul siia elama asunud. Ülekaetusest vabaneti põhiliselt küsitletavate aadresside täpsustamise järgus, kui vastavad isikud brutovalimist eemaldati. Arvestades, et viimaste aastate sisserändajad moodustavad kogurahvastikust väga väikese osa, ei saa selle mõju tulemuste esinduslikkusele olla märkimisväärne. Et uuringu eesmärk pole mitte agregeeritud hinnangute andmine, vaid eri käitumisvaldkondade seesmiste ja väliste seoste selgitamine, siis ei kujuta väikese hulga hilismigrantide valimist puudumine probleemi.

#### 3.2. Valimikava

Esinduslikkus üldkogumi suhtes tähendab rahvastiku regionaalse, demograafilise ja sotsiaalse koostise adekvaatset peegeldumist valimis. Väljavõtumäära arvutamiseks eristati üldkogumist kahe tunnuse, rahvuse ning vanuse alusel moodustatud kihte. Nimetatud kihistusskeem jaotas rahvastiku neljaks kihiks: nooremad eestlased, vanemad eestlased, nooremad mitte-eestlased ning vanemad mitte-eestlased. Kihtide defineerimisel eeldati, et väljalangevus kujuneb kihiti erinevaks; stratifitseerimise põhieesmärk oli neid kaoerisusi kompenseerida.

Konkreetsete väljavõtumäärade arvutamisel lähtuti eeldusest, et kadu rahvaloenduse andmebaasist koostatud valimist võib toimuda neljal erineval põhjusel: (1) surma, (2) väljarände, (3) küsitletava hetkeadressi mitteleidmise, (4) intervjuu eri asjaoludel tegemata jäämise tõttu. Et vältida vajadust töö käigus rahvaloendusandmestiku juurde tagasi pöörduda, olid alljärgvalt esitatud kaohinnangud etteteadvalt konservatiivsed.

(1) *Surm*. Suremushinnagutes lähtuti Eesti 1988-1989 aasta rahvuslõikelistest elutabelitest [Katus ja Puur 1992], eeldades, et rahvaloenduse ja küsitluse toimumise vaheajaks on täpselt viis aastat. Noorema naisterühma suremuse hindamisel lähtuti tabeltõenäosustest vanuste 15-45 ning vanema naisterühma puhul tabeltõenäosusest vanuste 45-65 jaoks. Hinnangu kohaselt moodustas kadu nii eestlaste kui mitte-eestlaste puhul nooremas vanusrühmas 0.5 protsenti, vanemas vanusrühmas aga vastavalt 5 protsenti.



(2) *Väljaränne*. Hinnangud põhinesid viimaste aastate andmetel Eesti väljarände kohta [Sakkeus 1991]. Nooremate eestlaste, vanemate eestlaste ja vanemate mitte-eestlaste puhul saadi rändest tingitud kaoprotsendiks 1, noorte mitte-eestlaste puhul 5 protsenti.

(3) *Mitteleidmine*. Hinnangud põhinesid *Saar Polli* varasematel kogemustel ja Sotsioloogiainstituudi poolt läbiviidud prooviküsitlusel. Aadressi mitteleidmisest tingitud kadu hinnati iga kihi puhul võrdselt 15 protsendi suuruseks.

(4) *Lõpetamata intervjuu*. Viimaste aastate küsitluskogemusest lähtudes eeldati, et küsitleva leidmisel õnnestub intervjuu teha 75 protsendi eestlaste ja 65 protsendi mitte-eestlastega. Silmas pidades kontakteerumiseks ning intervjuude tegemiseks ette nähtud pikemat ajavahemikku, võis eeldada, et lõpetatud küsitluste arv kujuneb tegelikkuses mõnevõrra suuremaks.

Summerides kolmel esimesena nimetatud põhjusel tekkinud väljalangevuse, osutus kao suurus valimikihti järgmiseks: noorematel eestlastel 17 protsenti, vanematel eestlastel, noorematel mitte-eestlastel ning vanematel mitte-eestlastel 21 protsenti. Viimatimainitud suuruste jagamisel eeldava intervjuu lõpetajate protsendiga saadi tulemuseks järgmine valimikihtide suurus:

Tabel 1. VALIMIKIHTIDE SUURUS

Valimikiht	Küsitletavate brutoarv	Küsitletavate netoarv	Valimimäär, %
Nooremad eestlased	2952	1802	1.53
Vanemad eestlased	2006	1182	1.56
Nooremad mitte-eestlased	2570	1315	1.87
Vanemad mitte-eestlased	1370	701	1.74
Kokku	8898	5000	1.65

Kihiti ebaühtlaste valimimäärade eesmärgiks oli tagada valimi täpsem vastavus üldkogumile, mitte teatud rahvastikurühmade normaalsest suurem esindatus.

### 3.3. Valimietapid

EPSU valimi moodustamine toimus kaheetapilise protseduurina: esiteks eraldati rahvaloenduse individuaalandmetest kihtvaliku teel 9000 isiku kirjed ning teiseks täpsustati nende aadressid.

*Esimene valimietapp.* Valimimoodustamise esimese sammuna eraldati rahvaloendusandmestikust kihtväljavõtu teel 9000 isiku kirjed. Arvestades eelpool esitatud arvutusi otsustati valimisse hõlmata 3000 nooremat eestlast, 2000 vanemat eestlast, 2600 nooremat mitte-eestlast ja 1400 vanemat mitte-eestlast. Otsene väljavõtt toimus SPSS-paketi RANDOM-protseduuri abil, väljavõtt teostati vastava rahvastikurühma 25-protsendilisse valikloendusse sattunute hulgast. Väljavõtualuse piiramine valikloendusse hõlmatutega oli

motiveeritud rikkalikuma taustinformatsiooniga. Koostiselt valikloendusse hõlmatud rahvastikuosa ülejäänust märgatavalt ei erinenud.

Järgnevalt esitati valimisse sattunute loenduskirjed küsitlusorganisatsioonile. Kuna loendatud isikute nimed ja aadressid arvutil olevast andmebaasist puudusid, tuli nende leidmiseks pöörduda Riigi Statistikaameti arhiivis säilitatavate loenduslehtede ja -nimekirjade poole. 106 küsitletava puhul ei õnnestunud arhiivist loendusaegset aadressi leida ning vastavad kirjed jäeti edasisest tegevusest kõrvale.

Pärast nimede ja aadresside väljakirjutamist loendusarhiivist täpsustati allesjäänud 8894 isiku küsitlusaegsed aadressid Keskaadressbüroos. Aadressbüroost leiti vajalik aadress 7676 inimese kohta. Vastavalt aadressbüroos talletatud informatsioonile jagunesid ülejäänud 1218 inimest kolme rühma: 320 olid pärast 1989 aastat Eestist lahkunud, 136 olid surnud ja 762 kohta aadressandmed puudusid. Surnuid ja Eestist lahkunuid võis käsitleda valimisaluse ülekaetusena ning nende kirjed kõrvaldati edasisest tööst. Need 762 inimest aga, kelle uut aadressi leida ei õnnestunud, jäid valimisse edasi.

*Teine valimietapp.* Esimese valimietapi lõpptulemusena jäi töösse 6914 isikut, kelle küsitlusaegne aadress oli keskaadressbüroost õnnestunud teada saada, ning 762 isikut, kelle puhul tuli piirduda loendusjärgse aadressiga. Esitatud jaotus ei tähenda muidugi, et kõik aadressbüroojärgsetel aadressidel ka elasid.

Esiolgu 5000 töösse antud küsitletava väljavalimiseks 6914 hulgast kasutati *SPSS*-tarkvara *RANDOM*-protseduurile tuginevat lihtjuhuvalikut. Pärast 3000 intervjuu tegemist, mil hakkas selguma valimikao suurus, edastati küsitlusorganisatsioonile täiendavad küsitletavakirjed. Kokku edastati täiendavaid kirjeid kolmel korral, vastavalt 227, 411 ning 650. Täiendavate kirjete valikul ei piiratud juhuvalikut nende küsitletavatega, kelle küsitlusaegne elukoht oli aadressbüroost teada.

Kokku edastati küsitlajakirjetele 6212 isiku nimed ja aadressid. Protseduurireeglite kohaselt pidid küsitlajakirjed respondendi elukohast mitteleidmisel püüdma hankida viimase asukoha kohta lisateavet erinevatest allikatest (omavalitsus, majavaldaja, naabrid jne). Kui mingit lisateavet saada ei õnnestunud, tuli kontrollida loendusjärgset aadressi. Kui töö käigus selgus, et küsitletav on vahepeal asunud mõnda teise Eesti piirkonda, anti ta üle mõnele vastavas piirkonnas tegutsevale intervjuerijale.

### **3.4. Küsitlustulemus**

Tabelis 2 on esitatatud kõigi töösse antud respondentide jaotus küsitlustulemuse seisukohalt. Tabelist selgub, et küsitlusorganisatsioonile edastatud 6212 juhtumist lõppesid intervjuuga 5021 ning mingil muul moel 1194. Viimane rühm sisaldab ka 50 niisugust küsitletavat, kellega mingil põhjusel kontakti võtta ei üritatudki.

Valimi netosuurust aluseks võttes kujunes tegelikuks vastamismääraks 0.93 protsenti. Iga küsitletu esindas seega üldkogumist 107 kaaskodanikku.

Tabel 2. KÜSITLUSTULEMUS TÖÖSSE ANTUD INTERVJUEERITAVATEL

Küsitlustulemus	Küsitletavate arv	Küsitletavate osakaal, %
Kokku töösse antud	6212	100.0
Küsitletud	5021 <sup>a</sup>	80.8
Mitte-küsitletud	1192	19.2
sh:		
üritati kontakteeruda	1142	18.4
üritatud kontakteeruda	50	0.8

<sup>a</sup> Küsitletute hulk sisaldab ka ühte eksikombel intervjuueeritud isikut, kelle nimi valiminimekirjas olnud respondentiga juhuslikult kokku langes.

### 3.5. Kaopõhjused

Toimumata võisid intervjuud jääda erinevatel asjaoludel. Osadel juhtudel oli tegemist aadresside täpsustamise käigus kõrvaldamata jäänud ülekaetusega, teistel juhtudel respondendi soovimatusena uuringus osaleda, kolmandatel ei saadud küsitletavaga lihtsalt ühendust jne. Ehkki igas olukorras oli kindlasti midagi unikaalset, on tabelis 3 püütud küsitlusetapil toimunud väljalangemist süstematiseerida, moodustades viis kaopõhjuste rühma.

Tabel 3. KAOPÕHJUSED

Kaopõhjus-	Küsitletavate arv	Küsitletavate osakaal, %
Surm	43	3.7
Raske haigus	117 <sup>a</sup>	10.2
Eestist lahkumine	188 <sup>b</sup>	16.5
Mitte-leidmine	380 <sup>c</sup>	33.3
Keeldumine	414	36.3
Kokku	1142	100.0

<sup>a</sup> Rühm sisaldab 19 alkohoolikut, 61 rasket haiget või invaliidi (parajasti haiglaravil viibinud, halvatud, kurtummad, vähi- või vaimuhaiged) ning 34 isikut, kes keeldusid vastamast tervislikel põhjustel.

<sup>b</sup> Rühm sisaldab 72 Eestist alaliselt emigreerunud ning 116 ajutiselt emigreerunud küsitletavat.

<sup>c</sup> Rühm sisaldab 2 vanglas viibinud küsitletavat.

Tabelist ilmneb, et kõige sagedasemateks kaopõhjusteks olid keeldumine ning respondendi mitteleidmine. Keeldumise all on siin mõeldud nõ lõplikku keeldumist. Kui küsitletava esmane reageering selge ja kategooriline: "Ärge minu poole enam pöörduge", "Mind ei huvita selles uuringus osalemine üldse" jne, siis taoliste inimestega (kokku 288) enam kontakti ei võetud. Kui keeldumine oli leebem: "Mul on kiire", "Mul pole praegu aega", "Ma pole kindel, kas tahan küsitluses osaleda", "Ma ei tunne end praegu hästi" jne, siis tehti katse veenda küsitletavat siiski uuringus osalema. Taolistel puhkudel saadeti respondentile täiendavate selgitustega kiri ning kui tarvis, saadeti tema juurde mõni kogenum intervjuerija. Kuigi üsna sageli andsid sedalaadi jõupingutused positiivse tulemuse, jäi 126 vähemotiveeritud keeldumist kokkuvõttes siiski muutmata.

Teistest kaopõhjustest väärrib tähelepanu küsitleva mitteleidmine, mille osatähtsus ületab väljarännet enam kui kahekordselt. Mitteleidmine hõlmab kõiki juhtumeid, kus küsitleva mingit ühendust ei saadud. Osadel juhtudel elasid inimesed ajutiselt mõnel teisel aadressil, mida intervjuerijal ei läinud korda teada saada (näiteks suvilas või maalviibimine). Küllalt sageli polnud vastajad aadressbüroost teadaoleval aadressil aga kunagi elanud, harvemini võis mitteleidmise taha peituda küsitlusest hoidumine ehk varjatud keeldumine. Mitteleidmise kõrge esinemissageduse põhjust tuleb näha eeskätt olemasoleva aadressiteabe ebatäpsuses ning nõukogudeajast pärinevas sissekirjutusel baseerivas registreerimissüsteemis, mis polnud orienteeritud inimeste tegeliku asukoha fikseerimisele.

### 3.6. Vastamiskordajad

Tabeli 2 põhjal saab arvutada vastamise üldkordaja, mille arväärtuseks on  $5021 : 6212 = 0.808$  ehk 80.8 protsenti. Üldkordaja lähtub lihtustatud eeldusest, et kõigil valimisse sattunud respondentidel oli põhimõtteline võimalus intervjueritud saada. Tegelikuses see eeldus täiel määral ei kehtinud, kuna küsitlevate hulgas oli ka inimesi, kes olid küsitluse ajaks surnud (45), selles osalemiseks liiga haiged (117) või Eestist lahkunud (188). Intervjuerimisrisk ei laiene ka neile 50le isikule, kellega kontakteeruda ei üritatudki. Käsitades surma, haiguse ning Eestist lahkumise tõttu väljalangemist kui ülekaetust, võib vastamiskordajat täpsustada arvestades riskiavatud respondente. Täpsustatud vastamiskordaja väärtuseks on  $5021 : 5815 = 0.863$  ehk 86.3 protsenti. Küsitlevate suhtumise iseloomustamiseks võib arvutada veel keeldumiskordaja, mille arväärtuseks on  $414 : 6212 = 0.067$  ehk 6.7 protsenti. Võrrelduna teise suurema viimasel ajal Eestis läbi viidud uuringu, Tööjõuuuringuga (n=9608), on EPSU kaokordaja mõnevõrra kõrgem. Parem vastamismäär on eeskätt seletatav uue maakondade statistikabüroodel põhineva küsitlusvõrgu respondentide leidmisel tehtud tõhusama tööga.

### 3.7. Valimi esinduslikkus

Valimi esinduslikkuse puhul on eesmärgiks süstemaatiliste erisuste puudumine valimi ja üldkogumi vahel, mis muidugi ei eelda kahe kogumi koostise üksühest vastavust. EPSU puhul on representatiivsuse hindamiseks kasutatud kahte tüüpi informatsiooni. Esiteks, vanuse, regiooni, asulatüübi, rahvuse, perekonnaseisu ning laste arvu kui analüüsi seisukohalt oluliste tunnuste lõikes arvutatud vastamis- ja kaokordajaid. Teiseks on esinduslikkuse hindamiseks kasutatud üldkogumi, vastajate ning mittevastajate struktuuri võrdlust samade tunnuste alusel.

Nagu valimikava koostamisel eeldatud, oli vastamiskordaja eestlastel kõrgem kui mitte-eestlastel. Lisaks Eestist lahkumisele seletub mitte-eestlaste madalam vastamistase sagedasema mitteleidmisega. Rühmadevaheline erisus jäi tegelikuses siiski oodatust väiksemaks ning tervikuna oli väljalangevus nii eestlaste kui mitte-eestlaste puhul koguni kümmekond protsendipunkti eeldatust madalam.

Linna- ja maaküsitlevate võrdlus näitab kõrgemat vastamistaset maal. Kaopõhjustest ilmneb, et väljalangemine oli maal madalam põhiliselt keeldumiste ning mitte-leidmiste harvema esinemise tõttu. Huvitav on märkida, et erisus linna- ja maarahvastiku vahel ilmnis üksnes eestlastel: mitte-eestlaste hulgas oli vastamiskordaja linnas 77.0 protsenti ja maal 75.6 protsenti; eestlaste puhul olid arvud vastavalt 81.2 ja 87.5 protsenti. Maakonniti ulatus

vastamiskordaja Järva-, Põlva-, Rapla- ja Viljandimaal üle 90 protsendi; suurima väljalangevusega Harju- ja Tartumaal jäi kordaja väärtus 75 protsendi piirimaile.

Tabel 4. VASTAMIS- JA KAOKORDAJAD KÜSITLETAVARÜHMITI

Küsitletavarühm - 1989 loendusel	Üld- vastamis- kordaja, %	Üld- kao- kordaja, %	Täpsustatud kao- kordaja, %	Keeldumis- kordaja, %	Mitte- leidmis- kordaja, %
Kokku	80.8	19.2	12.8	6.7	6.1
1969-1973	79.2	20.8	13.8	5.3	8.5
1964-1968	79.8	20.2	13.5	5.3	8.2
1959-1963	82.0	18.0	13.6	6.6	7.0
1954-1958	80.0	20.0	14.6	6.9	7.7
1949-1953	84.6	15.4	10.4	4.6	5.8
1944-1948	77.8	22.2	15.3	8.4	6.9
1939-1943	83.3	16.7	12.4	8.0	4.4
1934-1938	81.0	19.0	13.5	8.0	5.5
1929-1933	80.5	19.5	10.6	6.6	4.0
1924-1928	79.9	20.1	10.2	6.7	3.5
Eestlased	83.3	16.7	12.5	7.1	5.4
Mitte-eestlased	76.9	23.1	13.2	6.0	7.2
Linn	79.2	20.8	14.1	7.3	6.8
Maa	86.0	14.0	8.5	4.7	3.8
Harjumaa	76.4	23.6	16.5	8.3	8.2
Hiiumaa	80.5	19.5	17.1	9.8	7.3
Ida-Virumaa	81.5	18.5	9.5	4.6	4.9
Jõgevamaa	86.8	13.2	10.5	7.9	2.6
Järvamaa	92.0	8.0	4.9	4.3	0.6
Läänemaa	85.3	14.7	5.9	4.4	1.5
Lääne-Virumaa	85.4	14.6	11.1	5.0	6.1
Põlvamaa	91.6	8.4	4.2	2.5	1.7
Pärnumaa	83.3	16.7	12.0	6.5	5.5
Raplamaa	91.9	8.1	6.6	5.1	1.5
Saaremaa	84.4	15.6	7.8	4.3	3.5
Tartumaa	77.3	22.7	15.5	7.6	7.9
Valgamaa	82.8	17.2	11.2	6.7	4.5
Viljandimaa	91.8	8.2	6.1	3.9	2.2
Võrumaa	89.4	10.6	8.2	4.4	3.8
Kõrgharidus	83.2	16.8	13.1	6.5	6.6
Keskharidus	81.0	19.0	13.1	6.8	6.3
Põhi- või madalam	79.6	20.4	12.1	6.6	5.5
Abielus	81.8	18.2	12.2	6.3	5.9
Vallaline	77.7	22.3	14.3	7.4	6.9
Lesk	81.5	18.5	11.2	6.4	4.8
Lahutatud/lahuselav	80.4	19.6	14.1	7.4	6.7
0 elussündinud last	75.5	24.5	16.3	7.8	8.5
1	78.8	21.2	14.2	7.6	6.6
2	83.4	16.6	11.3	6.2	5.1
3	87.0	13.0	8.2	3.8	4.4
4+	87.3	12.7	7.4	4.9	2.5

Küsitletavate haridustasemega võib täheldada vastamiskordajaga positiivset seost. Väiksem kadu kõrgema haridustasemega küsitletavatel seletub nende harvema väljalangevusega suremuse ning haigestumuse tõttu, muud kaopõhjused hariduseti ei erinenud.

Tabel 5. ÜLDKOGUMI, KÜSITLETUTE JA MITTEKÜSITLETUTE VÕRDLUS

Küsitletavarühm 1989 loendusel	1989 loendus- rahvastik	EPSU küsitletud (n=5020)	EPSU mitteküsitletud (n=1192)
Kokku	9.8	9.2	10.2
1969-1973	9.8	9.2	10.2
1964-1968	9.4	9.9	10.5
1959-1963	11.0	10.1	9.3
1954-1958	11.4	9.9	10.4
1949-1953	10.9	9.6	7.4
1944-1948	9.1	9.6	11.6
1939-1943	9.4	10.9	9.1
1934-1938	9.8	10.6	10.5
1929-1933	9.7	10.0	10.2
1924-1928	9.6	10.1	10.7
Eestlased	59.8	62.7	52.9
Mitte-eestlased	40.2	37.3	47.1
Linn	74.8	75.0	82.9
Maa	25.2	25.0	17.1
Harjumaa	40.0	39.0	50.8
Hiiumaa	0.7	0.7	0.7
Ida-Virumaa	14.9	15.1	14.5
Jõgevamaa	2.5	2.6	1.7
Järvamaa	2.6	3.0	1.1
Läänemaa	2.0	2.3	1.7
Lääne-Virumaa	4.8	4.4	3.2
Põlvamaa	2.1	2.2	0.8
Pärnumaa	6.2	6.4	5.4
Raplamaa	2.3	2.5	0.9
Saaremaa	2.3	2.4	1.8
Tartumaa	10.5	10.1	12.5
Valgamaa	2.5	2.2	1.9
Viljandimaa	3.9	4.2	1.6
Võrumaa	2.7	2.8	1.4
Kõrgharidus	13.6	14.1	12.0
Keskharidus	54.2	53.8	53.4
Põhi- või madalam	32.2	32.0	34.6
Abielus	62.2	63.1	59.1
Vallaline	19.9	18.9	22.9
Lesk	7.0	7.1	6.8
Lahutatud/lahuselav	10.8	10.8	11.1
0 elussündinud last	24.2	22.6	30.8
1	27.1	26.6	30.0
2	34.9	35.0	29.3
3	9.7	10.9	6.9
4+	4.1	4.9	3.0

Et väljalangevus Eestist lahkumise ning mitteleidmise tõttu tasakaalustas nooremate vastajate seas vanemate põlvkondade sagedasemat haigestumust- ning suremuskadu, siis osutus vastamiskordaja vanusrühmiti üsna stabiilseks. Võrreldes omavahel erineva lapsesusega naisi, ilmnes vastamiskordaja nõrk positiivne seos laste arvuga. Kaopõhjuste struktuur lubab oletada, et nimetatud seose taga on enam lapsi omavate naiste madalam liikumus ning vahest ka lastega naiste suurem huvi uuringu temaatika vastu. Küsitlejate töökogemus vastamisinäitajates oluliselt ei kajastanud.

Teise viisina EPSU valimi esinduslikkuse hindamiseks on kasutatud küsitlusele vastanute ning mittevastanute koostise võrdlust üldkogumi, st 1989 aasta loendusrahvastikuga. Loendusandmete ja EPSU küsitletute võrdlusest ilmneb nende kahe kogumi tähelepanuväärselt suur sarnasus; vastanute selektiivsus üldkogumiga võrreldes ei ole ühegi vaadeldud tunnuse lõikes statistiliselt oluline. Üldjoontes kehtib sama ka mittevastanute kohta, kui mitte arvestada mitte-eestlaste ebarproportsionaalselt suurt arvu väljalangenute seas. Siinkohal on sobiv meenutada, et taolist tulemust arvestati juba valimikavas, kus mitte-eestlaste tarvis nähti ette kõrgem väljavõtumäär. Kuivõrd küsitletute rahvuskoostis on üldkogumile lähedane, siis ei peaks mitte-eestlaste suurem väljalangevus uuringu tulemusi mõjutama. Mittevastanute maakondliku jaotuse hälbed johtuvad madalamast väljalangevusest väiksemates maakondades. EPSU andmestiku analüüsi seisukohalt muudab üldkogumi ja küsitletute struktuuri sarnasus tarbetuks uuringutulemuste korrigeerimise järelkihistamise teel.

### 3.8. Kooskõla loenduskirjetega

Üheks võimaluseks uuringuandmete täpsuse hindamisel on nende kõrvutamise teistest allikatest pärineva informatsiooniga. Kui tavapäraselt on seda võimalik teostada agregeeritud andmete alusel, siis EPSU puhul tegi loendusandmestiku valimialusena kasutamine võimalikuks uuringust ning loendusest pärineva informatsiooni individijärgse ühildamise. Metodoloogiaülevaate koostamise raames on nimetatud kahest allikast pärinevate kirjete kooskõla võrreldud küsitletava sünniaja, perekonnaseisu, elussündinud ja elusolevate laste arvu ning majandustegevuse osas. Võrdluse teostamiseks on arvatud uuringuandmetele toetudes küsitletava seisund jaanuaris 1989 ning kõrvutatud seda vastava loendustunnuse väärtusega.

Tabel 6. LOENDUS- JA UURINGUTUNNUSTE VÕRDLUK, SEISUGA JAANUAR 1989

Tunnus	Loendus-, andmestik, %	Uuringu-, andmestik, %
Perekonnaseis		
Abielus	63.1	66.1
Vallaline	18.9	15.1
Lesk	7.1	6.7
Lahutatud/lahuselav	10.9	12.1
Majandustegevus		
Tegev	79.9	71.7
Mittetegev	21.1	28.3

Eri tunnuste puhul varieerus mittekokkulangevusmäär 3.0 protsendist sünniaasta puhul 11.8 protsendini perekonnaseisu ning 13.6 protsendini majandustegevuse puhul. Olles põhjustatud üksteist tasakaalustavatest juhuslikest eksimustest, ei ilmnenu sünniaja mittekokkulangevuses süstemaatilist hälvet. Suurema mittekokkulangevusega tunnuste puhul võis loendus- ja uuringuandmete süstemaatilist erinevust siiski täheldada. Nende illustreerimiseks esitab tabel 6 küsitlute jaotuse perekonnaseisu ning majandustegevuse järgi.

Tabelist nähtub loenduse ülehinnang vallaliste ja alahinnang abielus olijate osas. Loendus- ja uuringuandmete risttabelist selgub, et arvestatav osa uuringuandmete kohaselt vabaabielus olnud küsitlute oli loendusel registreeritud vallalistena. Teistes perekonnaseisurühmades on mittekokkulangevus olnud väiksem. Majandustegevuse osas seletub majandusaktiivsuse ülehinnang lapsehoolduspühkudel viibivate naiste erinevast liigitamisest. Loendusel arvati taolised naised töötavate hulka, uuringus on lapsehoolduspühkudel viibimist käsitletud mitteaktiivsuse ühe vormina. Loenduse mõningane alahinnang elussündinud laste osas näib johtuvat imikueas surnud laste mittetäielikust kajastatusest vanemate naiste loenduskirjetes. Kuna loendus- ja uuringuandmete suuremad mittekokkulangevused seletuvad definitsiooni- ja mõõtmiserisustega, siis ei ole see uuringuandmete analüüsil probleemiks. Pigem võib kirjete mittekokkulangemise sügavam analüüs pakkuda teavet loendusandmete sisulisemaks interpreteerimiseks.

## 4. KÜSITLUS

Nagu eelpool viidatud, oli nõukogudelaadse andmekorralduse üheks tunnusjooneks uuringute kaudu kogutava teabe alahindamine ning vastava andmekogumisviisi teisejärgulisus loenduse ning sündmusstatistikaga kui kahe teise põhilise andmekogumisviisiga võrreldes. Eestis suutsid üleminekuajal suurema hakanud nõudlusele uuringuteabe järele reageerida kõigepealt kommertsstruktuurid. Statistikaameti juures perebüzhetiandmeid kogunud, kuid vähesel professionaalse ettevalmistusega ning puuduliku juhendamisega intervjuerijate kaader osutus uuttes oludes vähesuutlikuks ning saadeti laiali. Riigi küsitlejavõrgu puudumisel otsustas EPSU Teadusnõukogu kasutada väliküsitluse läbiviimiseks kõige suutlikumat kommertsvõrku, kelleks prooviküsitluste tulemuste alusel osutus *Saar Poll*. Viimase puudulik valmisolek EPSU-laadse nõudliku töö läbiviimiseks tähendas Teadusnõukogu jaoks arvestatavat lisaülesannet, milleta poleks olnud mõeldav tarviliku andmekvaliteedi kindlustamine.

### 4.1. Küsitlejaõpe

Arvestades uuringu keerukust ning alalise töökogemusega professionaalsete küsitlejate nappust, pöörati uuringu ettevalmistamisel erilist tähelepanu küsitlejate valmisolekule. Lisaks *Saar Polli* poolt ettenähtud üldõppele läbisid kõik intervjuerijad 2-päevase ankeedikeskse õpetuse. Ajavahemikul detsember 1993 - jaanuar 1994 peeti uuringu ettevalmistuse raames kokku 5 õppeseminari asukohaga Tallinnas, Tartus ning Narvas. Eraldi voorud olid ette nähtud eesti ja vene keeles töötavatele intervjuerijatele. Tagamaks ühtlast korraldust viidi kõik seminarid läbi EKDK teadurite poolt, kes ühtlasi olid olnud küsimustiku eri osade väljatöötajateks.



Seminaridel pandi põhirõhk küsimustikus kasutatud mõistetele, andmedefinitsioonidele ning sündmusloolisest lähenemisviisist tulenevatele probleemidele ja sellega seotud intervjuerijatöö erinõuetele (meenutusvigade vältimine ning kooskõla tagamine eri ankeediosade vahel), samuti küsitlusprotseduurile. Erinevalt küsitlusorganisatsiooni tavapärasest praktikast, tegi iga küsitlajakandidaat seminari lõpul läbi individuaaltesti, mille kaudu selgitati tema töövalmidus. Test seisnes 1/2-1 tunni pikkuses proovintervjuus, kus vastaja rolli asus mõni EKDK teaduritest. Etteteatatult sisaldas iga sellise "küsitletava" elulugu mitmesuguseid vastuoksusi, mille märkamise oli üldise intervjuutehnika kõrval küsitlaja professionaalsuse hindamise põhikriteeriumiks. Ilmselt individuaalse lähenemisviisi uudsuse tõttu ei läbinud esimesel seminaril prooviintervjuud edukalt ükski paarikümnest kandidaadist. Järgnevatel kordadel liikus edukusnäitaja 50-70 protsendi piirimaile.

## 4.2. Küsitluskorraldus

Erinevate küsitlajakandidega tehtud prooviküsitluste üksmeelseks järeltöödeks oli välitööde protseduurilise külje vähene läbitöötatus, mis EPSU-laadse küsitluse korral jättis korraldajale vähe võimalusi probleemide ilmnemisel intervjuerijate tegevust suunata. Et selline suunamisvõimalus siiski toimiks, otsustati lisaks muudele abinõudele korraldada tööd ilma kohalikust koordinaatorist vahelülita küsitluskeskuse ning intervjuerija vahel.

Protseduuri kohaselt said prooviküsitluse edukalt läbinud küsitlajakandidaadid esimese tööülesandena enda tööpiirkonda kuuluvate küsitlajakandide nimekirjad. Nimekirjade koostamisel arvestati rahvaloendusandmestikust teadaolevat küsitlajakandide emakeelt: üksikute eranditega said eesti keeles töötavad intervjuerijad endale küsitlajakandideks eesti emakeelega, vene keeles töötavad küsitlajakandid aga venekeelsed respondendid. Nimekirjade koostamisel püüti jälgida põhimõtet, et küsitlajakand ja küsitlajakand üksteist ei tunneks. Potentsiaalselt võis tekkida olukord, kus intervjuerija ja intervjueritav osutusid üheks ja samaks isikuks (EPSU puhul tekkis üks niisugune kokkusattumus).

Nimekirjadega tutvumise järel tuli igal küsitlajakandil koostada kogu tööperioodi hõlmav individuaalne kontakteerumisgraafik. Graafiku eesmärgiks oli ühendada keskne küsitlusprotsessi kontroll intervjuerijate suhtelise autonoomiaga sellisel viisil, mis minimeeriks edaspidise infovahetuse küsitlajakandide ning keskuse vahel. Küsitlusgraafikute alusel saadeti keskusest küsitlajakandidele kontaktkirjad. Lisaks teabele küsitluse kohta sisaldas kontaktkiri konkreetsele respondendile määratud intervjuerija aadressi ning tema telefoninumbri. Ümbrikusse lisatud vastuskaardil paluti küsitlajakandil määrata enda jaoks sobiv kontaktivõtaviis, milleks võis olla oma telefoninumbri teatamine, ise küsitlajakandile helistamine või kohe sobiva intervjuuaja väljapakkumine. Vastuskaardil oli respondendil võimalus anda teavet ka oma keele-eelistuse kohta. Kui küsitlajakand keele-eelistus osutus eeldatust erinevaks, määrati talle tarvidusel uus intervjuerija.

Hilisema analüüsi kohaselt saatis vastuskaardi tagasi 2894 ehk 57.6 protsenti küsitlajakandidest. Kui maa- ja linnaelanike hulgas oli kaardisaatjate osakaal peaaegu ühesugune, siis põlisrahvastiku puhul oli vastanuid 8 protsendi võrra keskmisest enam. Vastuskaardi saatsid keskmisest sagedamini tagasi ka vanemad ning kõrgema haridustasemega küsitlajakandid. Vähesele variatsioonile vaatamata osutus kaardi saatmine või mittesaatmine suhteliselt heaks indikaatoriks respondendi suhtumise ja võimalike eelseisvate küsitlusraskuste kohta. Silmatorkavalt kõrge kaardisaatnute osakaal jaanuaris-veebruaries lubab oletada küsitluse algul

meediakanalites läbiviidud tutvustuskampaania, mis hõlmas korduvaid artikleid kesk- ja kohalikus pressis, samuti spetsiaalseid tele- ja raadiosaateid, tõhusust.

Vastuskaardi saatnute hulgas oli eelistatuim kontaktiviis intervjuerijale enda telefoninumbri teatamine (41.3 protsenti), millele järgnes konkreetse intervjuuaja väljapakkumine (35.4 protsenti) ja ise küsitlajale helistamine (19.7 protsenti). Ülejäänud 3.5 protsendil juhtudest sisaldas vastuskaart teavet mujale elama asunud küsitletava uue aadressi kohta. Arvestades eeltoodut aitasid kontaktkiri ja vastuskaart lisaks vastaja informeerimisele vältida tühikäike ning alandada sellega küsitluskulusid.

Vastuskaardi saamist pidi linnaoludes töötav küsitlaja ootama 10 päeva, selle mittelaekumise korral pidi küsitletav ise respondendiga ühendust võtma. Tühisõitude vähendamiseks oli maaoludes ette nähtud 20-päevane ooteaeg, misjärel küsitletavale saadeti korduskiri. Kui ka sellele vastust ei saanud, tuli küsitlajal responndi juurde sõita.

Kui küsitletav kontaktkirjale vastas või avaldas mingil muul moel valmisolekut uuringus osaleda, siis leppis intervjuerija sobiva küsitlusaja ning -koha respondendiga kokku. Ankeetide täitmiseks kasutati intervjuud, välja arvatud üksikud erandjuhtumid, kus intervjuust keeldunud küsitletav avaldas soovi küsimustiku isetäitmiseks. Kolmest kirjeldatud viisil laekunud ankeedist lülitati kaks hiljem andmebaasi, üks tuli aga puuduliku täidetuse tõttu tööst kõrvale jätta. Õnnestunud kontaktivõttudest ning kokkulepitud intervjuudest informeeris küsitlaja regulaarselt küsitluskeskust (enamasti toimus infovahetus telefonitsi). Omamaks ülevaadet üldisest töökulust kanti teave kontakteerumiste ja laekunud ankeetide kohta keskuses valimifaili viisil, mis lubas seda võrrelda esialgsete kontakteerumisgraafikutega.

Kui ilmnes, et küsitletav oli Eestist lahkunud või surnud, siis kontakteerumiskatsed arusaadavalt lõpetati. Samamoodi toimiti ka juhtudel, kui respondent osutus institutsionaliseerituks (haiglaravil või, kahel juhtumil, vanglas olevaks), või kui küsitletav ei soovinud intervjuud anda tervislikele põhjustele viidates. Keeldumistena käsitleti juhte, kus küsitletav oma mitteosalemissoovist kirjalikult või suuliselt teatas. Nende küsitletavatega, kelle keeldumismotiive võis pidada ajutiseks, üritati veelkordselt ühendust võtta. Mitmetel juhtudel aitasid täiendavad selgitused uuringu eesmärkide kohta veenda taolisi respondente siiski osalema.

Kõigi juhtudel dokumenteeris küsitlaja kontakteerumise üksikasjad (telefonikõnede ning kohalkäikude aja ja tulemuse, sugulastelt, naabritelt, majavaldajalt, kohalikest omavalitsusest ning mujalt saadud lisainformatsiooni jms), samuti oma tegevuse lõpptulemuse spetsiaalsele protokollile. Kui intervjuud ei toimunud ja küsitlaja leidis olevat ammandanud kõik enda käsutuses olevad võimalused intervjuu teostamiseks, saatis ta täidetud protokollid küsitluskeskusele. Keskuses kontrolliti väljalangemise motiveeritust ning tarvidusel suunati see koos asjakohaste soovitustega intervjuerijale tagasi. Lõplik otsus väljalangemise aktsepteerimise või tagasilükkamise osas tehti kahepoolselt küsitlusorganisatsiooni ja EPSU Teadusnõukogu poolt.

### **4.3. Küsitlajate arv**

Andmekvaliteedi tagamise eesmärgil oli küsitlajate töölevõtmisel nõudeks vähemalt 50 intervjuu sooritamise. Nimetatud arvu tuleb käsitleda kui kompromissi kvaliteedinõuete ning

küsitlajate arvu vähendamise tingitud transpordikulude suurenemise vahel. Küsitlajate koguarvuks oli esialgselt planeeritud 70.

Kuna enamus *Saar Polli* küsitlajatest töötas osaajaga ning ei suutnud oma esialgselt kohustusest kinni pidada, siis tuli välitöödel tegelikult kasutada 93 intervjuerijat, kellest 65 töötas peamiselt eesti ning 28 peamiselt vene keeles. Lõpptulemusena piirdus vähema kui 50 intervjuuga läbiviimisega 36 küsitlajat, 39 küsitlajat tegid 50-75 ja 15 üle 75 intervjuu. Kolm küsitlajat katkestas töö, olles sooritanud vähem kui 10 intervjuud; nendele määratud respondendid tuli jagada teistele küsitlajatele. Maksimaalne kontaktivõttude arv ühe küsitlaja kohta oli 294, millest ankeediga lõppes 234.

Tabelis 7 esitatud andmetest ilmneb otsene seos küsitlaja poolt tehtud intervjuude arvu ning töö kvaliteedi vahel: mida enam intervjuusid, seda paremad tulemused. Huvitava seigana väärib veel mainimist uute intervjuerijate paremad tulemused võrreldes nende *Saar Polli* küsitlajatega, kes omasid varasemaid kogemusi arvamus- ning muude lihtsakoelise mate sotsioloogiliste küsitlustega.

Tabel 7. KÜSITLUSKÄIK KÜSITLJARÜHMITI

Küsitluskulg	Küsitlaja poolt sooritatud intervjuude arv			
	<50	50-99	100+	Kokku
Intervjuu kestus, %				
Alla 30 min	0.2	0.6	0.0	0.4
30-59 min	7.6	9.5	21.7	11.0
60-89 min	38.2	41.5	59.3	43.6
90-119 min	33.0	34.5	16.2	31.3
120+ min	21.1	13.9	2.7	13.7
Keskmine kestus, min	94.5	89.1	72.2	87.5
Küsitlitava suhtumine uuringusse, %				
Väga huvitatud	19.8	24.5	22.7	23.2
Mõnevõrra huvitatud	55.5	56.5	65.5	57.7
Mõnevõrra mitte huvitatud	22.1	17.3	10.4	17.2
Täiesti mitte huvitatud	2.7	1.8	1.4	1.9
Vastuskaot üldkordaja, %	15.5	20.6	13.6	18.5
Kuupäevakao levimus, %				
Aasta	6.1	5.2	2.8	4.9
Kuu	22.4	18.7	12.8	18.9

#### 4.4. Küsitluskuu

Esialgselt oli EPSU küsitlusperioodiks plaanitud ajavahemik jaanuarist maini 1994. Töö käigus selgus, et mitmesuguste asjaolude tõttu pole võimalik esialgselt graafikust kinni pidada, ning tegelikult viidi põhiosa intervjuudest läbi ajavahemikul jaanuarist augustini, üksikud intervjuud toimusid aga veel septembris-novembris 1994. Arvestades EPSU tagasivaatelist lähenemisviisi, ei olnud küsitlusperioodi pikenedes mõju tulemuste kvaliteedile.

Tabel 8. KÜSITLUSKUU

Kuu	Intervjuude arv	Kumulatiivprotsent
Jaanuar	128	2.5
Veebruar	837	19.2
Märts	1044	40.0
Aprill	880	57.5
Mai	861	74.7
Juuni	566	86.0
Juuli	296	91.8
August	264	97.1
September	93	99.0
Oktoober	32	99.6
November	20	100.0

#### 4.5. Küsitluspaik

Enamus (73.5 protsenti) intervjuudest toimus küsitletava kodus. Ülejäänutel oli küsitluspaigaks töökoht (10.1 protsenti) või mõni muu paik: küsitlaja kodu, tema töökoht, mõni avalik koht jms (kokku 16.0 protsenti). Küsitluskoha varieeruvus respondendirühmiti oli vähene, ootuspäraselt on vanemate küsitletavate puhul töökoht keskmisega võrreldes tagasihoidlikumalt esindatud.

88.2 protsenti küsitletavate puhul toimus intervjuu kõrvaliste isikute juuresolekuta, 8.2 protsendil juhtudest viibis küsitluse juures mõni respondendi lastest, 2.4 protsendil juhtudel abikaasa, samuti 2.4 protsendil juhtudest mõni muu täiskasvanu. Kõrvaliste täiskasvanute vähene juuresolek annab alust oletada, et küsitlussituatsioonil tulemustele märgatavat mõju ei olnud.

#### 4.6. Küsitluskestus

Tulenevalt sündmusloolisest lähenemisviisist ning talletatava informatsiooni hulgast, kujunes intervjuu kestus suhteliselt pikaks. Mõõdetuna täidetud tabeliveergude arvuga kujunes fikseeritud elusündmuste keskmiseks arvaks 17-18, alates põlvkonnast 1959-1963 sündmuste arv küsitletava vanuse kasvades oluliselt ei suurenenud. Kirjapandud üksiksündmuste arv oli suurim elukohavahetusi puudutavas ning väikseim kooselusid käsitlevas ankeediasas.

Küsimustiku ettevalmistamisel eeldati intervjuu keskmiseks pikkuseks poolteist tundi, tegelikkuses oli keskmine küsitluskestus 87.5 minutit. Kolm neljandikku intervjuudest mahtus kestusvahemikku tunnist kuni kahe tunnini. 30 intervjuu puhul tuli küsitlus mingitel asjaoludel katkestada ning teisel korral lõpule viia.

Vastajarühmast sõltuvalt esinesid küsitluskestuses süstemaatilised erisused. Noorimas põlvkonnas kujunes intervjuu keskmiseks pikkuseks 75.7 minutit, vanimas põlvkonnas ulatus küsitlusaeg aga 94.2 minutini; enam kui kahetunnise pikkusega intervjuude arv oli vanemate naiste puhul üle kolme korra keskmisest kõrgem. Hariduseti kulges eralduspiir keskhariidusega ning keskhariiduseta küsitletavate vahel: keskhariiduse puudumine tähendas intervjuule keskmiselt 10 minutist lisaaega. Võrreldes küsitlusperioodi algul tehtud intervjuude kestust

lõpuperioodi omadega ilmneb, et töökogemuse lisandudes langes küsitluseks tarvilik aeg 100 minutilt (jaanuar-veebruar) 80 minutile (juuni-september). Nagu eelpool mainitud, avaldas kestusele tugevat mõju ka konkreetse küsitlaja töökogemus antud uuringus. Küsitlajatel, kes sooritasid kogusummas alla 50 intervjuu kulus ühe intervjuu tegemiseks 94.5 minutit, üle 100 intervjuu sooritanud küsitlajatel kulus seevastu keskmiselt ainult 72.2 minutit. Enam kui poolteist tundi kestnud intervjuude osakaal ulatus esimese rühmas 54.1 protsendini, teises rühmas moodustasid pikad intervjuud napilt 18.9 protsenti.

Lääneriikides väljakujunenud arusaamade seisukohalt suhteliselt pikk EPSU küsitluskestus oli enne välitööde algust Teadusnõukogus tõsiseks arutlusaineks. Eriti rõhutasid sellega seonduvaid võimalikke ohte väliseksperdid. Tabelis 10 esitatud erineva kestusega intervjuude võrdlusest selgub, et pikema intervjuukestusega ei kaasnenud siiski halvenemist andmekvaliteeti iseloomustavates näitajates. Selgesti on aga andmetes jälgitav seos küsitluskestuse ning küsitlaja poolt täheldatud meeldetuletusraskuste vahel.

Tabel 9. KÜSITLUSKÄIK KÜSITLUSKESTUSETI

Küsitluskulg	Küsitluskestus, minutit				Kokku
	-59	60-89	90-119	120+	
Küsitluse sujuvus, %					
Väga hea	56.6	41.4	35.9	39.2	41.1
Hea	36.4	47.4	48.4	38.3	45.5
Mõningate raskustega	6.1	9.7	12.2	15.6	10.9
Suurte raskustega	0.9	1.5	3.5	6.9	2.5
Korduskontakteerumine küsitletavaga, %	9.9	11.2	12.2	12.7	11.6
Loendus- ja uuringuandnestiku lahknevusindeks <sup>a</sup>	113.1	99.7	100.4	95.7	100.0

<sup>a</sup> Indeks põhineb loendus- ja ankeedikirjete summaarsest lahknevusel, indeksi moodustamisel on kasutatud tunnustena sünniaega, perekonnaseisu, elussündinud ja elusolevate laste arvu, haridustaset ning majandustegevust 1989 jaanuaris.

#### 4.7. Küsitlus kvaliteet: intervjuerijahinnangud

Iga intervjuu lõppedes tuli ka küsitljal endal vastata paarile küsimusele, mis puudutasid respondendi üldist suhtumist ning intervjuu käigus ilmnenu probleeme. Ehkki kirjandusest pole leida palju tõendusi küsitlajahinnangute ning tegeliku andmekvaliteedi sõltuvusest (tulemuste tõepärasuse mõttes), esitame need tulemused aset leidnud suhtluse iseloomustamiseks.

Küsitlajahinnangutest selgub, et 23.2 protsenti intervjueritavatest ilmutas uuringu vastu suurt huvi, 57.7 protsendil on täheldatud mõningat huvi ning 17.2 protsenti on küsitlajad iseloomustanud kui huvivaeseid. Täielikku huvipuudust uuringu vastu on märgitud vaid harvadel juhtudel (1.9 protsenti). Respondendirühmiti on keskmisest suuremat huvi tähele pandud põlisrahvastiku ning kõrgema haridusega küsitletavate hulgas. Vanemate vastajate suhtumine on olnud vaid pisut vähem aktiivne kui noorematel. Vastupidiselt mõningatele ootustele ei kajastu küsitlajahinnangutes seos laste arvu ning vastaja suhtumise vahel.

Vastaja suhtumisest mõnevõrra erinevad on hinnangud intervjuu sujuvusele. Küsitlejate hinnangul kulges 41.1 protsenti intervjuudest väga hästi ja 45.5 protsenti hästi. Mõningaid raskusi on nimetatud 10.9 protsendil, suuri raskusi aga 2.5 protsendi intervjuude puhul. Et intervjuerimisraskused esinevad sagedamini vanemate põlvkondade hulgas, siis on põhjust oletada, et enamikul juhtudel on raskused seisnenud elusündmuste toimumisaja meenutamises. Üldhinnanguna tuleb intervjuerijate kogemust pidada siiski väga positiivseks, isegi vanima kohordi juures (sündinud 1924-1928) on 78 protsenti intervjuusid hinnatud hästi või väga hästi sujunuks, mõningaid raskusi on täheldatud 17 ning suuri raskusi vaid 4.9 protsendil. Süstemaatilist variatsiooni intervjuu sujuvuses võib lisaks vanusele täheldada veel seoses küsitleva haridustasemega.

Ka hinnangute osas lubab küsitlejarühmade võrdlemine korrata järeldust intervjuerimiskogemuse olulisusest: mida enam intervjuusid küsitlejad on teinud, seda paremini hindavad nad enda koostööd respondentiga, ning seda harvemini võib kohata viiteid küsitlusraskustele.

## **5. KODEERIMINE, SISESTUS JA ANDMEPUHASTUS**

Soovimatute järeparanduste vältimiseks oli EPSU andmete kodeerimine ja sisestus küsitlejavõrgust lahutatud ning toimus Teadusnõukogu juhendamisel selleks otstarbeks loodud eraldi üksuses. Andmete sisestusele eelnes keerukama väärtusvaruga tunnuste: elukoha, ameti, hariduse (eriala), rahvuse ja kodukeele kodeerimine. Ameti kodeerimisel kasutati rahvusvahelisele ametiklassifikaatorile ISCO-88 baseeruvat liigitust, ametid kodeeriti kahekohalise kooditäpsusega. Haridustunnuse kodeering jälgis rahvusvahelise haridusklassifikaatori ISCED liigitust kahekohalise kooditäpsusega. Rahvuse ja keele kodeerimisel kasutati vastavat loendusliigitust. Küsitluses esinevad sünni- ja elukohad kodeeriti omavalitsuse (linna/valla) täpsusega.

Et kodeerimine puudutas ajavahemikku 1880 aastatest tänaseni, raskendasid sünni- ja elukohakoodide puhul tööd haldusjaotuse korduvad muutused Eestis. Võrreldavuse huvides viidi küsitletavate elu- ja sünnikohad kaasaagele (1989 aasta loenduse) alusele, mis tähendas kaartide, kohanimeregistrite ja mitmesuguste muude allikate kasutamist. Hariduse ja ameti osas komplitseeris kodeerimist asjaolu, et Eestis pole rahvusvaheliste klassifikaatorite juurutamine lõpule viidud ning puuduvad vastavad käsiraamatud. EPSU puhul tuli spetsiaalselt uuringu tarbeks koostada ajutised kodeerimisjuhendid. Parema tulemuse saavutamiseks oli kodeerimine sisestusega ühendatud: mõlemad tööoperatsioonid teostati sisestusrühma poolt, kes lõpptulemusena vastutas nõuetekohase andmebaasi formeerumise eest.

Erinevalt levinud praktikast ei näinud EPSU töökorraldus ette sisestuseelset ankeetide täidetuskontrolli, mis tavapäraselt teostatakse küsitlusjuhendajate poolt. Andmepuhtuse esmakontroll oli siin ühendatud andmesisestusega; ISSA tarkvarapakatile tuginev sisestusprogramm sisaldas hulga loogilisi- ja vahemikkontrolle, mis operaatorit vigadest informeerisid. Kui ankeedis sisalduvaid vastuolusid polnud võimalik eri küsimuste võrdlemise teel kõrvaldada, siis sisestus peatati ning vastav ankeet saadeti küsitlejale täpsustamiseks tagasi. Välistamaks meelevälgeldusi parandusi, saadeti intervjuerijale mitte kogu ankeet, vaid üksnes selle täpsustust vajavatest lehekülgedest tehtud koopiad. Küsitlejatel tuli taolistel puhkudel respondentiga uuesti kontakteeruda ning vastuolud kõrvaldada.

Vähendamaks vigade esinemissagedust ning küsitletavate mitmekordset tülitamist, teostas EKDK küsitluse algul ankeetides esinenud vigade analüüsi ning informeeris ringkirja kaudu kõiki küsitlejaid tüüpilisematest eksimustest. Vaatamata tehtud jõupingutustele tuli respondente siiski teistkordselt tülitada 11.6 protsendil juhtudest (582), küsitlusjärgselt teostatud analüüsist ei ilmnenud vigade esinemissageduse süstemaatilisi erisusi ei vastaja- ega küsitljarühmiti. Viie ankeedi puhul osutusid sisestuse käigus ilmnenud puudused nii ulatuslikuks, et need tuli tööst üldse kõrvale jätta; nimetatutest nelja juhul oli tegemist vaimsete puuete tõttu institutsionaliseeritud küsitletavatega, ühel juhtumil aga isetäidetud ankeediga.

Pärast andmesisestust ning esmakontrolli formeeriti failikirjetest terviklik andmebaas, mis teisendati *SPSS-Windows* formaati ning allutati põhjalikule kontrolltöötlusele; kokku testiti andmete kooskõlasust mitmesaja erineva seose suhtes. Igal korral, kui selgusid vastuolud, kõrvutati failikirjeid küsitlusarhviivis hoitava ankeediga. Enamikul juhtudel osutus taolise võrdluse tulemusena võimalikuks vastuolud kõrvaldada. Valdava enamiku parandatud vigade allikaks oli olnud küsitlaja tähelepanematus ning suutmatus küsimustiku eri osades antud vastuseid loogilisse seosesse viia, sisestajavigade osakaal ulatus umbes ühe kümnendikuni. Ehkki küsimustiku koostamisel oli nn lahtiste vastusvariantide kasutamise vajadust minimiseeritud, tuli andmepuhastuse käigus osa lahtistest vastusvariantidest läbi vaadata ning ümber kodeerida. Andmepuhastuse lõpul teostatud esialgse ning puhastatud andmefaili võrdluse kohaselt ulatus teostatud ühiktäpsustuste arv ligikaudu 20 tuhandeni, puhastuse käigus tehtud kontrollimised puudutasid kolme neljandikku küsimustike koguhulgast. Mõnevõrra ootamatult ei toonud analüüs välja sõltuvusi parandatud vigade arvu ning küsitlaja- või küsitlavarühmade vahel.

## 5.1. Küsimusjärgne vastuskadu

Käsitletud üldkaole lisaks sõltub andmebaasi kvaliteet ning analüüsiks kasutatav teabehulk üksikküsimuste mittevastamusest. Ehkki konkreetne kaomäär küsimuseti teatud määral varieerub, iseloomustab EPSU andmebaasi tervikuna küsimusjärgse kao väga madal tase. Enamiku küsimuste puhul on tegemist üksikute mittevastamisjuhtudega, mille mõju analüüsitulemustele on minimaalne. Reeglina johtus mittevastamine konkreetse küsimuse puhul ilmnenud meenutusraskustest; vastamisest keeldumist esines päris üksikutel juhtudel. On tähelepanuväärne, et ankeedi kõige intiimsemate küsimuste puhul leidis vaid kolm repondenti, kes ei olnud nõus nimetama oma möödunud nelja nädala seksuaalvahekordade arvu. Ühe jaoks neist kolmest ületas taoliste üksikasjade pärimine taluvuspiiri ning küsitlajat katkestas intervjuu. Respondentide avatus on kooskõlas EPSU spetsiaalse abortide puudutanud prooviküsitluse tulemustega ning kinnitab, et Eestis räägitakse kõnealuselt valdkonnast märksa vabamalt kui mitmetes lääneriikides [Anderson, Katus, Puur, Silver 1993]. Põlvkonniti suhteliselt stabiilne ning perioodandmestikuga kooskõlas olev suhe abortide ja sündide arvu vahel (2:1) lubab väita, et EPSU andmestik kajastab reproduktiivvaldkonda küllalt suure tõepärasusega.

EPSU-laadse sündmusloolise andmestiku puhul on andmekvaliteedi seisukohalt kriitilise tähtsusega kuupäevaandmete täielikkus. Sündmuste toimumisaja kuulise täpsusega meenutamine oli vastajatele kindlasti üheks tõsisemaks probleemiks: oli ju vahel tegemist enam kui poolsajandi eest aset leidnud sündmustega. Arvestades küsimuste detailsust võib igati

heaks lugeda tabelis 10 esitatud tulemust, mille kohaselt enam kui 95 protsenti respondentidest suutis meenutada kõiki küsitud aastaarve. Väärrib rõhutamist, et üheski küsitleva enda elukäiku puudutas moodulis ei ulatunud sündmusaastat mitteteadnud respondentide osakaal üle 0.4 protsendi.

Tabel 10. KUUPÄEVAKADU KÜSITLETURÜHMITI, ELULOOLOSED PÕHIMOODULID

Moodul/Küsitleturühm	Aasta teadmata, %	Kuu teadmata, % <sup>a</sup>
Kogu valim/Kõik moodulid	4.9	19.9
Elukaaslaste lugu	0.2	0.8
Sünnilugu	0.1	0.5
Raseduslugu	0.4	2.1
Vanematekodu	4.4	17.1
Rändelugu	0.0	0.9
Leibkonnalugu	0.1	0.8
Haridus- ja töölugu	0.1	0.6
Eestlased	3.9	16.4
Mitte-eestlased	7.6	26.1
Linn	5.5	20.6
Maa	4.5	18.4
Harjumaa	5.6	18.8
Hiiumaa	0.0	0.0
Ida-Virumaa	8.2	32.6
Jõgevamaa	7.1	29.9
Järvamaa	1.3	4.0
Läänemaa	15.8	41.2
Lääne-Virumaa	3.2	18.9
Põlvamaa	9.0	19.0
Pärnumaa	6.3	25.4
Raplamaa	4.6	8.1
Saaremaa	1.7	18.5
Tartumaa	1.4	1.1
Valgamaa	1.9	5.8
Viljandimaa	3.1	23.4
Võrumaa	0.0	4.4
Kõrgharidus	1.4	14.3
Keskharidus	3.9	16.0
Põhi- või madalam	10.5	32.3
Abielus	2.4	10.5
Vallaline	4.8	18.8
Lesk	9.0	29.9
Lahutatud/lahuselav	6.8	24.8
0 elussündinud last	2.3	11.6
1	5.7	19.4
2	5.6	22.1
3	5.4	21.4
4+	7.3	25.7

<sup>a</sup> Ei kajasta olukordi, kus respondent ei suutnud meenutada täpset sündmuskuud, kuid mäletas, millisel aastajal sündmus aset leidis.

Mis puutub sündmuskuusse, siis oli selle unustamine ootuspäraselt sagedasem. Siiski ei ulatunud siingi mittevastanute osakaal üheski põhimoodulis peale rasedusloo (2.1 protsenti) üle ühe protsendi. Meenutamiskeskuste põhihulk näib valdavalt pärinevat vanematekodu käsitlevast küsimustikuosast, kus peamiseks probleemiks on olnud oma varases lapseas



surnud õdede-vendade sünniaja meenutamine. Uuringuterviku seiskohalt puudub sellel andmestikuosal õnneks kriitiline tähendus.

Tabelis 10 on muu hulgas esitatud kuupäevakao erisused vastajarühmiti. Kahe- kuni kolmekordseid kaoerisusi võib täheldada vanuseti, hariduseti ning põlisuseti, kuupäevakao vanuserisustest on üksikasjalikumalt juttu allpool. Kogenumad küsitlajad paistavad olevat osanud respondendi meenutamispüüdlusi toetada, saavutades vähemkogenutega võrreldes üle kahe korra madalama mittevastamistaseme. Kuupäevakao üldiselt madala taseme tõttu ei peaks kirjeldatud erisustel siiski analüüsitulemustele märgatavat mõju olema.

## 5.2. Küsitluskvaliteet vanemates põlvkondades

EPSU Teadusnõukogu otsus tõsta küsitluskogumi ülemist vanuspiiri 49 eluaastalt 69 eluaastani kaasas paratamatult teatud riski halvendada andmekvaliteeti ning alandada uuringuandmete usaldusväärsust. Tuli ju vanematel küsitletavatel noorematega võrreldes meenutada palju kaugemaid elusündmusi ning teatud vanusest alates võis see ealiste iseärasuste tõttu ülejökäivaks osutada. Lisaks võis vanemate põlvkondade puhul eeldada kõrgemat kaomäära halvenenud tervise ning küsitluses puudutatud teemade diskreetsuse tõttu. Ehkki prooviküsitluste resultaadid olid julgustavad, polnud enne põhiküsitlust päris selge, mil määral vanemate põlvkondade eripära andmekvaliteedis kajastub.

Nagu eelpool juba mainitud, osutusid valimikao vanuserisused erilaadiliste kaopõhjuste vastastikuse tasakaalustumise tõttu minimaalseteks. Vanemates põlvkondades eeldatud suuremat kadu haiguste ning suremuse tõttu kompenseeris mitteleidmise, väljarände ja keeldumiste harvem esinemine. Et EPSUs käsitlemist leidnud probleemid olid vanematele naistele arusaadavalt veidi vähem aktuaalsed kui noorematele, siis ei mõju ootamatuna esimeste pisut väiksem huvi uuringu vastu. Kuigi vanemate naiste puhul läks uuskontaktumist tarvis sagedamini kui noorte juures, muutuvad põlvkondlikud erisused elusündmuste arvu arvestamisel minimaalseks.

Kooskõla loenduskirjetega oli vanematel põlvkondandel koguni parem. Loendus- ja uuringuandmestiku kooskõlaindeksi vanussõltuvus viitab pigem indeksi arvutamisel aluseks olnud sündmuste (eriti perekonnaseisu ning majandusliku aktiivsuse) kui andmestike endi mittekooskõla vanussõltuvusele. Suurem tõenäosus mittekooskõla tekkimiseks on ju eeskätt nendes rühmades, kus inimeste seisund parajasti muutub, seega vanusskaala alumises ning ülemises osas. Spetsiifiliselt vanematele naistele iseloomulikuks võib pidada vahest üksnes sagedasemat erisust elussündinud laste arvus: nimelt ei kajastunud loendusandmestikus osa nende imikueas surnud lapsi. EPSU andmestiku seisukohalt tuleb muidugi tähele panna, et tegemist on vaadeldaval juhul loendus-, mitte uuringuandmete ebatäpsusega.

Ainus valdkond, kus vanematelt naistelt noorematega võrreldes ebatäpsem informatsioon laekus, oli vanematekodu. Vanemate naiste juures täheldatud meenutamiskasused olid koondunud küsimusse, millega selgitati küsitletava õdede-vendade arvu ning nende sünniaegu. Suuremad raskused oma õdede-vendade sünnikuupäevade meenutamisel ei seostu siin ilmselt mitte niivõrd küsitletute endi vanuse, kuivõrd kõrgema sündimustaseme ning eriti kõrgema imikusuremusega nende põlvkonnas.

Tabel 11. KÜSITLUSKVALITEET PÕLVKONNITI

Küsitluskvaliteet	Sünnipõlvkond										
	1924 1928	1929 1933	1934 1938	1939 1943	1944 1948	1949 1953	1954 1958	1959 1963	1964 1968	1969 1973	Kok- ku
Vastamiskordaja	79.9	80.5	81.0	83.3	77.8	84.6	80.0	82.0	79.8	79.2	80.8
Kaokordaja	10.2	10.6	13.5	12.4	15.3	10.4	14.6	13.6	13.5	13.8	12.8
Suhtumine uuringusse, %											
Väga huvitatud	22.4	21.9	21.7	22.4	22.8	23.2	27.6	24.3	21.1	24.8	23.2
Mõnevõrra											
...huvitatud	55.2	54.6	57.3	58.4	60.2	57.0	54.1	60.6	60.7	59.6	57.7
...ebahuvitatud	20.0	20.3	17.4	17.9	15.5	19.2	16.1	14.1	16.4	14.3	17.2
Täiesti ebahuvitatud	2.4	3.2	3.6	1.3	1.5	0.6	2.2	1.0	1.8	1.3	1.9
Aastakadu, %											
Kõik moodulid	11.8	7.9	7.4	6.2	4.8	4.5	3.5	3.4	0.9	0.1	4.9
Kooselud	0.2	0.0	0.4	0.2	0.2	0.0	0.0	0.2	0.4	0.0	0.2
Sünnilugu	0.2	0.2	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Raseduslugu	1.2	0.2	0.6	0.6	0.2	0.4	0.2	0.2	0.0	0.0	0.4
Vanematekodu	10.1	7.7	6.6	5.2	4.4	4.1	3.2	2.8	0.4	0.1	4.4
Rändelugu	0.0	0.2	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Leibkonnalugu	0.4	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Haridus- ja töölugu	0.2	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.0	0.2	0.1
Kuukadu, %											
Kõik moodulid	33.9	31.3	30.8	26.0	19.9	18.1	14.7	10.3	6.3	6.0	18.9
Kooselud	1.7	0.9	1.6	1.2	0.7	1.1	0.4	0.7	0.4	0.0	0.8
Sünnilugu	1.5	1.5	0.8	0.8	0.0	0.2	0.2	0.2	0.0	0.0	0.5
Raseduslugu	2.2	3.7	3.8	3.6	3.1	1.8	2.4	1.5	1.3	0.2	2.1
Vanematekodu	31.0	29.0	28.0	23.9	18.2	15.3	12.7	9.5	5.7	6.0	17.1
Rändelugu	1.0	2.9	1.7	1.2	1.4	0.7	1.0	0.0	0.0	0.2	0.9
Leibkonnalugu	2.7	1.9	1.2	1.2	0.0	0.2	0.2	0.6	0.2	0.0	0.8
Haridus- ja töölugu	1.9	0.9	0.6	0.4	0.7	0.7	0.4	0.2	0.2	0.2	0.6
Korduskontakteerumine, %											
	12.8	12.6	12.2	10.8	11.5	10.6	10.3	13.5	11.8	10.4	11.6
Loendus- ja uuringuandmestiku lahknevusindeks											
	106.4	100.1	84.1	80.6	70.8	74.0	85.8	93.4	131.9	178.4	100.0

### 5.3. Andmefail

EPSU andmepanka hoitakse dokumenteeritud ja märgenditega varustatud *SPSS-Windows* süsteemfaili kujul. Failiga on liidetavad respondentide kirjed 1989 aasta loendusandmestikust. Andmefaili maht ilma tuletatud tunnusteta on 10 MB, muutujate arv on 2063.

Arvestades puuduva kuupäevateabe väga vähest esinemissagedust, otsustati vastavasisulised lüngad põhiliste elusündmuste osas (kooselud, sünni- ja raseduslugu, rände- ja eluasemelugu, haridus- ja töölugu) programm-menetluse abil täita. Vanematekodu ning leibkonnaloo osas, kus puuduva teabe osakaal oli kõrgem, programm-menetlust ei rakendatud. Protseduuri käigus asendati kõigepealt kuuandmete kvartalitähised vastava kvartali keskmise kuuga; juhtudel, kus sündmuse toimumiskvartal polnud aga teada, kasutati kuuväärtuse saamiseks *SPSS-Windowsi* UNIFORM-protseduuri. Üksikute puudevate aastaväärtuste kõrvaldamiseks kasutati ekspertkorrigeerimist. Kuupäevakao väiksuse tõttu ei ole teostatud muudatustel

analüüsitulemustele mingit praktilist mõju, küll aga võimaldavad need oluliselt lihtsustada andmebaasi töötlust.

## VIITED

- Anderson, Barbara, Kalev Katus and Brian Silver (1994). Developments and Prospects for Population Statistics in Countries of the Former Soviet Union. *Population Index*, Vol.60, No.1, pp.4-20.
- Anderson, Barbara, Kalev Katus, Allan Puur and Brian Silver (1993). Characteristics of Women Having Abortions in Estonia. *Proceedings of the 22nd International Population Conference*. Montreal-Liège, IUSSP, Vol.1, pp.215-234.
- Anderson, Barbara, Kalev Katus, Allan Puur and Brian Silver (1994). The Validity of Survey Responses on Abortion: Evidence from Estonia. *Demography*, Vol.31, No.1, pp.115-132.
- Blossfeld, Hans-Peter, Alfred Hamerle and Karl Ulrich Mayer (1989). *Event History Analysis. Statistical Theory and Application in the Social Sciences*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, Publishers.
- Darsky, Leonid (1986). *Evolucija rozhdaemosti v SSSR po dannym obschesojuznyh vyborochnyh obsledovanii*. Paper presented to Deuxieme Seminaire Franco-Sovietique de Demographie, September 15-19. Suzdal.
- Katus, Kalev (1990). Demographic trends in Estonia throughout the centuries. *Yearbook of Population Research in Finland*. Helsinki, pp.50-66.
- Katus, Kalev (1991). *Fertility Trend and Regional Differences in Estonia*. RU, Series B, No.12. Tallinn, EKDK.
- Katus, Kalev (1994). Fertility transition in Estonia, Latvia and Lithuania. *Wolfgang Lutz, Sergei Scherbov and Andrei Volkov (Eds). Demographic Trends and Patterns in the Soviet Union Before 1991*. London and New York, IIASA, Routledge, pp.89-111.
- Katus, Kalev and Allan Puur (1992). *Eesti rahvastiku suremustrend elutabelite analüüsi põhjal*. RU, Series B, No.14. Tallinn, EKDK.
- Katus, Kalev and Allan Puur (1993). *The 1989 Estonian Population and Housing Census: Data Description*. RU, Series A, No.33. Tallinn, EKDK.
- Katus, Kalev, Allan Puur, Luule Sakkeus and Brian Silver (1995). *Eesti Pere- ja Sündimusuuring. Metodoloogiaülevaade*. RU, Series A, No.39. Tallinn, EKDK.
- Mitmekeelne Demograafiasõnastik. Eesti väljaanne*. (1993). Tallinn, EDA.
- Pavelson, Marje and Dmitri Mihhailov (1982). Sredovaya aktivnost' gorodskikh semei: na puti k territorial'noi obschnosti. *Mikk Titma (Red). Sotsiologitskeskie issledovanija v Sovetskoi Pribaltike*. Vilnius, Institut Filosofii, Sotsiologii i Pravo AN LSSR, s. 264-280.

- Sakkeus, Luule (1991). *Post-War Migration Trends in Estonia*. RU, Series B, No.15. Tallinn, EKDK.
- Sakkeus, Luule (1994). The Baltic States. *Solon Ardittis (Ed). The Politics of East-West Migration*. New York, St. Martin's Press, pp.68-85.
- Titma, Mikk (1993). *Zhiznennyie puti odnogo pokolenia*. Moskva, Nauka.
- Tuma Nancy, Michael Hannan (1984). *Social Dynamics: Models and Methods*. New York, Academic Press.
- UN ECE (1992a). *Fertility and Family Studies in the Countries of the ECE Region*. A Project Undertaken by the Population Activities Unit of the Economic Commission for Europe with Financial Support from the United Nations Population Fund. Geneva, United Nations.
- UN ECE (1992b). *Fertility and Family Studies in the Countries of the ECE Region*. Questionnaire and Codebook. New York, United Nations.
- UN ECE (1993). *Fertility and Family Studies in the Countries of the ECE Region*. Standard Recode Files and Standard Country Reports. Geneva, United Nations.

# INTRODUCTION

## 1. EFFS PROJECT OUTLINE

Estonian Family and Fertility Survey (EFFS) is a national project of the 1990 round of Family and Fertility Surveys in Europe, coordinated by the Population Activities Unit of the UN ECE (UN ECE PAU). The planning of the EFFS was initiated by the Estonian Demographic Association and Estonian Interuniversity Population Research Centre (EKDK) in 1991. Being the first survey of its kind in Estonia, besides its regular tasks EFFS had to contribute significantly to the reorganisation of post-Soviet statistical system and developing the integration of population data into decision-making process. In its turn to secure consistent results, the EFFS arrangements frequently turned out to be much broader in scope than procedures of a regular survey. From another perspective, the participation in the international cooperation grants the Project an advantage of conceptual and methodological consistency with modern population science, being a good impulse for the advancement of demography in Estonia. It also provides excellent possibilities to compare Estonia with other European countries, thus helping to locate the country in the international context of population development.

### 1.1. Survey Environment

The Soviet totalitarian system, prevailing in Estonia for half a century, implied among other features the respective statistical organisation. This peculiar and Moscow-centred organisation has been effective until recently [see Anderson, Katus, Silver 1994]. In order to provide understanding of arrangements for survey-based data collection in general, and activities connected with the EFFS, four characteristics of the societal environment and/or its statistical consequences should be underlined.

The first characteristic of the survey environment has been the absence of relevant sampling routines. Although a series of statistical surveys on family and fertility matters had been carried out in the former Soviet Union covering all republics [Darsky 1986], samples were restrictively produced by the Central Statistical Office in Moscow. As a result, neither sample frame nor the knowledge of sampling was developed in national statistical institutions. Though the description of sampling procedures for the above-mentioned surveys have never been published in details, it was apparent that the samples of all-union surveys were clearly non-representative for the populations of smaller republics like Estonia, even if they were representative to the total population of the Soviet Union. Due to small sample size and the use of multistage clustering procedures, in some cases only a few urban and/or rural districts were covered. Estonian Statistical Office, according to archive materials, never raised the question of achieving the representativeness at the national level.

On the other hand, under Soviet statistical organisation the scientific community had only restricted access to the population data, and definitely no access to individual-level data. Even if having the excellent theoretical knowledge of sampling procedures, it could not be applied for survey planning in practice. In Estonia, the theoretical knowledge of sampling which had

been developed apart from applications by statisticians at Tartu University, proved to be out of date and non-applicable when relevant basis for frame building became available. Thus, instead of application of existing routines, the development of the sample frame for nationwide representative surveys and elaboration of respective procedures turned into innovative part of EFFS.

The second characteristic of the survey environment to be noted was the lack of the professional interviewers' network. To some extent this feature stems from the previous one: as proper sample frame and sampling procedures for nationwide surveys were not developed in Estonia, professional interviewers' network covering the whole country was out of need. During the Soviet period the Statistical Office operated a kind of the prototype network to collect data on family expenditures, however, lacking regular professional training and supervision, it became non-competitive and was dissolved at the beginning of transition period in Estonia.

Apart from statistical surveys, various sociological studies carried out by research units as well as other institutions, used occasional networks established on temporary basis. Interviewers who had worked on occasional basis and were potentially available for EFFS, lacked systematic training and were relatively unfamiliar even with the basic interview techniques, not speaking about the qualifications required for the collection of life-history data on sensitive matters. Therefore, in the preparations of EFFS, the training and management of the interviewers' network became one of the central issues. Despite extremely time-consuming initiative, the interviewer training and close supervision of fieldwork by EFFS research team was considered unavoidable to ensure the data quality.

The third feature of the survey environment, confronted in the planning and implementation of EFFS, was the absence of appropriate project-based funding practices in Estonia. Regarding surveys, this problem reflects partly the general under-estimation of population information for decision-making procedures, particularly the under-estimation of survey data which was considered irrelevant for governmental agencies. By statistical institutions, the interest was limited to the aggregate tabulations, no individual-level data were acknowledged as a constituent of official statistics. In these circumstances, the financing a major survey like EFFS from state budget appeared to be entirely new idea. As the budget planning was built restrictedly on the annual basis, there were no rules to apply funding for the entire project period. Rapid inflation preceding to the currency reform and restraints typical to transitional economy did not make the task easier. Therefore, the EFFS preparations demanded a lot of clarification efforts and lobbying with officials responsible for budget formulation. For example, the EFFS as a separate item was several times discussed in the Budget Commission of the Parliament, and was voted for three times in the full assembly of the Parliament. Finally it resulted in the allocation of one third of the needed resources, making the parallel international fund-raising efforts essential.

The fourth aspect of the survey environment is related to comparability and consistency of existing and continuously produced population data. Regarding surveys in Estonia, there had been several sociological studies accomplished by research institutions during the Soviet period.

Most of them have been small-scale and concentrating on some specific problem or subgroup of population, however, in some cases they have been rather extensive (for example the longitudinal study of youth cohorts by Institute of Sociology [Titma 1993] and complex

surveys of Tallinn population by Urban Research Institute [Pavelson, Mihhailov 1982]). These surveys typically drew their samples on enterprise/organisation employee' or electors' lists with questionable coverage and virtually no background information. As nearly none of them had properly documented the methodology, these surveys cannot provide the consistent and comparative results despite the datasets themselves have been preserved at the computing centre of Estonian Radio. Thus, data of all these existing surveys stand apart and cannot be linked neither to each other nor to the administrative datasets.

To avoid such perspective, EFFS has made conscious efforts to secure conceptual and definitional linkage with other major datasets. Despite met with suspicion by statistical institutions, it has to date succeeded with the 1989 census, Estonian Labour Force Survey and annual databases on births and deaths allowing for comparability on aggregate as well as on individual level.

In short, features characteristic to the post-Soviet survey environment added to the preparation and implementation of EFFS several tasks which are normally not covered in the framework of one particular survey.

## **1.2. Management and Organisation**

Given the above-described circumstances it became evident that in order to prepare and implement EFFS, efforts of all institutions and interested individual researchers should be united. After discussions at the Estonian Demographic Association it was decided to establish the Working Group for planning and management of EFFS, and Estonian Interuniversity Population Research Centre was delivered the implementation functions of the survey.

The EFFS Working Group which combined researchers from academic institutions and officials from various governmental agencies was officially established in May 1992. Invitations to join the Working Group were sent to all relevant institutions which representative(s) formed the membership of the Working Group: Hans Hansen (Tallinn University), Kalev Katus (Head, EKDK), Aili Kelam (Institute of Sociology), Toivo Kitvel (Tallinn University), Jüri Kruusvall (Tallinn University), Jüri Kõre (Tartu University), Dmitri Mihhailov (Russian Cultural Society), Katrin Paadam (Tallinn Technical University), Allan Puur (EKDK), Asta Põldma (Secretary, Estonian Demographic Association), Leena Rõbakova (Secretary until 1993, Estonian Demographic Association), Luule Sakkeus (EKDK), Helbe Sinimäe (Tartu University), Mati Sundja (Statistics Estonia), Lee Tammemäe (Estonian Society of Perinatology), Lembit Tepp (Statistics Estonia), Andres Vikat (EKDK), Rein Võormann (Institute of Sociology). Later Hill Kulu (Tartu University) replaced Jüri Kõre in the Working Group. In position of foreign experts, Barbara A. Anderson (University of Michigan), Brian D. Silver (Michigan State University) and Jan M. Hoem (Stockholm University) were invited to participate in the Working Group.

The EFFS Working Group took the responsibility for all stages of the survey planning and implementation including the coordination of data analysis. Aside substantive aspects, the Working Group disposed also the financial resources allocated for the survey. To be independent from medium-level administrative bureaucracy, the EFFS Working Group has been operating as a separate unit under the Governmental Commission for Population Statistics.



Contiguous implementation functions of the survey were performed by the Estonian Interuniversity Population Research Centre. During 1992-1993 this body, under the guidance of the EFFS Working Group, elaborated the methodological approach and the survey instrument, prepared the sample frame, drew the sample, organised the FFS-specific interviewer training, supervised the fieldwork, organised data entry and accomplished data cleaning.

The survey questionnaire and instrument was intensively discussed in the EFFS Working Group and pretested in three pilot studies. In late 1993 the sample was drawn from the latest census, and verified for current addresses. Based on the results of the pretests, the Working Group selected interviewers' network of *Saar Poll* survey agency to carry out the fieldwork. The interviews were done in the period through January to November 1994. In 1995 the data entry and cleaning have been completed. Currently, the Methodological Report and Standard Tabulations are prepared as the first two basic publications to be followed by the Country Report. In next year the survey is expected to approach the stage of analysis.

Due to underdeveloped official institutions and the conditions of transitional society, the EFFS Working Group had to carry out broad scale of activities and responsibilities normally handled by regular structures. Despite being time-extensive, participation in these activities was considered inevitable to achieve the goals of the Project. Particularly the data quality has been secured by the participation of the EFFS Working Group in close supervision of fieldwork activities and data entry. Experience gained during the EFFS undertakings have been later discussed in the Governmental Commission, and partly already been used for planning other activities in the field of population statistics. The Labour Force Survey which followed closely the organisation of the EFFS, and building up the state interviewers' network on the basis of county statistical bureaux could be mentioned as examples.

### **1.3. Programme**

Methodologically the EFFS builds on the life course approach which implies the focus on timing, sequencing and spacing of various events in the lives of individuals. In modern population science this approach have proved to be useful to integrate individual behaviours and macro-level trends in demographic processes [Tuma, Hannan 1984; Blossfeld et al. 1989 etc]. In the EFFS, information on events is collected retrospectively and covers the time span between birth and moment of the interview. Programme of the EFFS is basically consistent with the UN ECE core questionnaire involving event histories on parental home, partnerships and marriages, conceptions and births, educational, occupational as well as housing and migration careers [UN ECE 1992b]. In addition to core questionnaire, some modules were noticeably expanded to follow the demographic development in Estonia providing an integrated insight into the population reproduction during the whole postwar period, and establish an adequate database for population-related policies in the country.

The first module with noticeably expanded programme is concerned with wanted and unwanted pregnancies and family planning perceptions, i.e the general context of pregnancy outcomes. Estonia has been characterised by very high pregnancy and abortion rates on the background of low fertility; induced abortion rate had exceeded the corresponding birth rates for several decades [Anderson, Katus, Puur, Silver 1993]. The EFFS is expected to highlight the social and individual circumstances which bring about different decisions with respect to

childbirth and abortion. As most of the first pregnancies are conceived in cohabitation, the union formation, particularly the first partnership becomes an issue of high relevance. In respect of timing, the rejuvenation trend of menstruation, sexual relations and union formation was accompanied by earlier childbearing. However, the end of the 1970s marked a reversal of the trend in respect of childbearing. Currently all the mentioned processes undergo complex change, introducing the discontinuity into the development of Estonian population.

Under the modern type of reproduction, the level of population replacement is decisively determined by the prevalence of third births. In Estonia during 1968-1990 the proportion of women who gave birth to the third (and higher order) child was sufficient to secure the population replacement level among the native-born population, neither earlier (since the end of the 1920s) nor later this has been the case. The parity-specific approach to matrimonial, sexual, abortion and fertility histories is expected to outline the more and less probable pathways which lead to the three-child family.

The second module with expanded programme concerns the migration and residential career. During the post-war period the formation of Estonian population has greatly been affected by the immigration from the regions of the former Soviet Union: currently more than 26 per cent of the population of Estonia are foreign-born, together with the second generation of the immigrants foreign-born population constitutes more than one third [Sakkeus 1991; 1994]. Compared to the native-born population, the immigrants are characterized by a distinct pattern of demographic behaviour as well as younger age composition. In immigration trend two major waves demarcating the principal enlargement of migration hinterland which lies underneath the sharp differentiation of Estonian immigrants.

Aside the massive segment of foreign-born population, Estonia has also been characterized by an intensive internal migration, including forced migration during the 1940-1950s. Internal migration has caused shifts in the spatial distribution of population bringing about specific socio-spatial regions, respectively pulling and pushing the population. The evaluation of migrations from the viewpoint of internationally comparable definitions of short- and long-distance moves as well as short- and long-term moves which has not been possible from other sources, must be regarded as a special merit of the EFFE. The analysis of relationship between timing of migrations and family-life events aims to bring out the multidimensional links between spatial mobility and family careers.

Aside the discussed, somewhat broader approach compared to UN ECE core questionnaire was applied also in modules concerning parental home and household welfare. On the other hand, the societal context supported the exclusion of some questions of the core questionnaire which appeared irrelevant due to rapid transformation. Under transitional, and therefore unstable conditions it proved to be of little use to attempt to receive perpetual information on value perceptions. Some other attitudinal questions could not be included because of incomparability across the social context. However, in general, according to UN ECE 1994 implementation record, the Estonian questionnaire had the third highest rank of comparability among 12 countries.

## 1.4. Survey Population

The EFFS sampling plan followed the UN ECE recommendations for national surveys. The target population constituted of females born in 1924-1973, i.e. of 20-69 years old in 1994, within this range all permanent residents were eligible. The selection of cases was decided to be without geographic clustering, reflecting fully the social composition of population including institutionalised persons (e.g. prisoners, nuns etc). The determination of the sampling plan involved both statistical and cost considerations. The original sampling plan considered two alternative approaches, which corresponded to different assumptions about the financing of the Project. Under the *minimum* plan, 5000 females aged 20-69 were to be interviewed. Under the *maximum* plan, 7500 females plus 2500 males were to be interviewed. All of the additional 2500 female cases under the maximum plan would have involved the women in fertile age of 20-49. In fact, the *minimum* plan was the one that was used in conducting the survey.

Regarding the representativeness, on the whole, the distribution of EFFS respondents approximates closely the eligible population. Particularly, it is notable that the completion rate is the same for the older cohorts as for the younger ones. The EFFS data does not require any additional weights or adjustments to results to make them more representative of the target population.

To draw the sample, the microdata of the 1989 population census were used as a sample frame. Given the aim of EFFS, the sampling unit was an individual. To date the 1989 census data is the only dataset ensuring the comprehensive coverage of Estonian population; besides providing a list of individuals, census data also supplied each case with a diverse background information. The design of the sample began already in 1991, when the Estonian Interuniversity Demographic Research Center, in cooperation with the Russian and Estonian Statistical Offices and researchers affiliated with the Population Studies Center of the University of Michigan worked together to transfer microdata from the 1989 census into PC-format. Restructuring and documentation of the dataset were prerequisites to the use of the census for a sampling frame [Katus, Puur 1993].

Although the EFFS sampling plan followed the UN ECE recommendations, in its coverage EFFS departed to some extent from the standard approach. Namely, the age range of the target population was extended above the limit of 50 years, i.e. above the fertile age. There has been two basic motivations of adaptation such an extended definition of target population. Firstly, it was motivated by the absence of earlier family and fertility surveys in Estonia or any other sources of relevant information. Only possible way to obtain the data on life histories for older cohorts was their incorporation to the survey sample. Secondly, the extended age range allowed to cover important features and turning points in the Estonian postwar demographic development, particularly in fertility trend.

Unlike other countries, in Estonia fertility was characterised by the lack of postwar baby-boom, and the rise of period fertility and maintaining the replacement level from the end of 1960s [Katus 1994]. This unprecedented trend on the European context could be analysed only having the relevant data on older female cohorts as well. On the other hand, in the post-war period, particularly during the first post-war decade, the formation of Estonian population was strongly affected by an extensive immigration from the territories of the former Soviet Union. The immigrant population was and still is characterized by a different pattern of demographic

behaviour with the impact of heterogeneity on the population development in Estonia being notably great. In this regard, having data on all immigrant cohorts by time of arrival would enable to address the formation of immigrant population, thus benefiting the understanding of adaptation and integration processes in Estonia. As already mentioned, due to the efforts made throughout the data collection stage, the inclusion of older cohorts involved no losses in data quality.

## **1.5. Funding**

The funding for the EFFS has come from multilateral sources. Most of the direct costs for the design and fieldwork have been covered from the Estonian budget sources. Despite the allowance from the 1993 budget covered only one third of the required funds for the Project, additional funding on annual basis during the following years has increased the budget contribution notably. The budget financing was managed via the Statistical Office, and supervised by the Governmental Commission of Population Statistics.

Particularly during the initial stages of the survey preparations, the input by the Estonian Interuniversity Population Research Centre has been critical. EKDK supported the EFFS with the restructuring of the 1989 census data, related expertise and work environment, and, most valuably, time of the researchers, throughout all stages of the implementation. In several occasions EKDK directly covered the survey expenses and took the risks of pre-financing the EFFS undertakings.

Foreign experts of the EFFS, Barbara Anderson and Brian Silver, supported the inclusion of specific modules in the survey instrument from their research funds. These researchers also funded the special pilot survey on abortion conducted in 1992 (NICHD Grant Nos. HD-19915 and P30 HD-10003). Foreign experts Anderson, Silver and Hoem volunteered their time as well as covered their expenses while participating in the Working Group sessions.

At the concluding stage of implementation, the United Nations Population Fund (UNFPA) supported the completion of the EFFS fieldwork, particularly data entry and cleaning as well as preparation of standard tabulations and methodological report (Grant No. EST/95/P01). UNFPA grant has proved essential to solve the financial problems accumulated by the end of the fieldwork. The first analyses of the EFFS data, including the analytical part of the methodological report and standard tabulations, have been funded by the Estonian Science Foundation (Grant No 1485).

As a national part of the 1990 round of Family and Fertility Surveys, EFFS is coordinated by the Population Activities Unit of UN ECE. The design and methodology of the EFFS has been based on guidelines worked out by UN ECE [1992a; 1992b; 1993]. UN ECE has also provided EFFS with technical advice, materials and training. The coordination by UN ECE PAU will be continued during the analytical stages of the EFFS.

## **2. SURVEY INSTRUMENT**

### **2.1. Definitions of events**

Estonian Family and Fertility Survey was aimed to collect data on central events and circumstances in the lives of individuals, particularly on parental home, socialisation environment, educational career, family, sexual, abortive, contraceptive and fertility histories, work, residential and housing career. Definitions of events underlying various domains of life histories build on internationally recommended statistical definitions. The application of these definitions to FFS-data collections was achieved through collaborative arrangements between the countries participating in the European project, the result of this work is reflected in the UN ECE PAU core questionnaire [UN ECE 1992b]. Concepts related to the presentation of EFFE data are available from another material [UN ECE 1993]. It is noteworthy that applied definitions are comparable to those used internationally in population censuses and vital statistics.

While the definitions of events were agreed between countries, preparation of the instructions for their application remained the task of individual countries. To ensure the comparability between populations undergoing different stages of demographic development and having diverse socio-economic conditions, the emphasis of these instructions depended indeed on specific circumstances in each country. As the development of statistical methodology has started in Estonia only recently, the elaboration of survey instructions turned, to significant extent, into innovative and responsible activity. Conducting a survey parallelly in two languages necessitated the materials to be prepared in Estonian and Russian languages. Both versions considered the specific features of demographic development between native and non-native populations. The questionnaire and the accompanying instructions have been included in the appendices of this volume.

### **2.2. Questionnaire Design**

Generally, the design of the EFFE questionnaire relies on the core questionnaire of the European family and fertility survey meant to record detailed event histories on leaving parental home, partnerships, births, educational and occupational careers [UN ECE 1992b].

As mentioned above, considering the empirical context, regarding childbearing, pregnancy, migration and residential histories as well as parental home and household welfare, the programme of the survey was extended. In the UN ECE core questionnaire, all life histories were to be recorded from January of the year the respondent became 15 years old. In the EFFE questionnaire, the record of events starts from age 14. The lower age limit was required because for older cohorts compulsory education did not go beyond that age.

Altogether, the questionnaire included 13 modules, listed as follows:

1. Household characteristics
2. Partnership history
3. Current partnership
4. Maternity history and children
5. Other pregnancies
6. Fertility regulation
7. Attitudes toward childbearing
8. Other values and beliefs
9. Parental home
10. Migration history
11. Living arrangements history
12. Education and work history
13. Current housing conditions

From the viewpoint of interview technique, including complex skip patterns and alternative wordings, the questionnaire appeared quite demanding, requiring flexible reaction to trace the lines of respondent's life history. Using table-format, different life histories were recorded in separate sections of the questionnaire. When the number of events exceeded the columns of the tables, additional sheets were attached.

To ensure the consistency between the scheduling of different events the interviewer filled parallelly a special summary life history chart. In this chart, each row corresponded to one calendar year and separate columns were foreseen for recording different events. When filling in a certain event in the questionnaire, the interviewer also put a mark in the chart. The purpose of the chart was to provide an interviewer with a tool for quick verification of the consistency of the interview. Using the table, she could detect and correct the potential inconsistencies before leaving the respondent. In addition to the questionnaire the EFFS survey instrument included detailed interviewer and procedural instructions, response cards, interview protocols and contact letters.

The ethnic composition target population of EFFS made it necessary to prepare the survey instrument parallelly in two languages. The original version was developed in Estonian, however, the Russian version was prepared simultaneously. All questions in the Estonian and Russian variants are intended to be identical; in some cases, the translation of the survey into Russian led to the identification of improvements needed in the Estonian version. The English variant is translated from the Estonian version of the questionnaire.

### **2.3. Pretests**

To evaluate the survey methodology, instrument, field procedures and potential survey organisations three pilot surveys were conducted. The first pilot survey took place in Tallinn (the capital of Estonia) and in Viljandi city and the rural areas of Viljandi County in February-March 1993. In Tallinn the interviews were conducted by *Saar Poll* survey agency Tallinn (30 respondents), in Viljandi the pilot was performed by the interviewers from County Statistics Bureau (18 respondents in Viljandi and 20 in rural communities). The main aim of the first pilot was to test an Estonian version of questionnaire, another goal was to test for differences

in response rates in urban and rural settings. Oversampling of older women was meant to provide an insight into the prevalence of recall difficulties across various modules.

The first pretest established a basis for revising the survey protocol, it also established a need for extensive training of interviewers to collect the life history data. On the basis of the experience obtained from the first pilot study, a revised version of the survey instrument was developed simultaneously in Estonian and Russian languages which served as the basis for the second pilot study. The new version was evaluated and further elaborated at the EFFS expert seminar in Tallinn, May 1993.

The second pilot study was conducted in July-August 1993. The aim of the second pilot was to verify the feasibility of using the 1989 census as the sampling frame and to test the survey instrument for the Russian-speaking respondents. For this reason the pilot study was carried out in the city of Narva -- located in northeast Estonia, bordering on Russia, and with a population that is 97 per cent non-Estonians, who were also then living in a politically strained situation (a referendum was underway on the question of separation of the region from Estonia). This was also a region that was most likely to have been impacted by emigration. In the second pilot another survey organisation was tested (Department of Sociology of the Institute of Philosophy, Law and Sociology of the Estonian Academy of Sciences). The test showed that the census addresses of persons could be updated with the help of the central address bureau in Tallinn. More importantly, the pretest indicated that conducting a survey among the Russian-speaking population need not be of special concern. None the 62 respondents who were contacted refused to be interviewed and only 3 were not located due to changes of address.

In addition, as part of the preparatory work, a special pilot survey on abortion was conducted in Tallinn in 1992. This survey helped to demonstrate that the population of Estonia was much more likely to respond accurately about some aspects of their reproductive behavior than is generally true in fertility surveys. For example, more than 80 per cent of the respondents who had had an abortion in 1991 reported having such an abortion, compared to more typical rates of 30-50 per cent found in surveys in the United States [Anderson, Katus, Puur, and Silver 1993, 1994].

On the basis of the pretests, the final version of the survey instrument was prepared by October 1993. The Working Group made a decision to use *Saar Poll* survey organisation for the fieldwork as the only survey organisation in the country at the time which had a nationwide interviewer network suitable for carrying out the survey.

### **3. SAMPLING PROCEDURE**

#### **3.1. Target Population and Sample Frame**

The target population of EFFS consisted of women aged 20-69 in the beginning of 1994. Within the mentioned age-range all permanent female residents were eligible, no specific categories of population (non-natives, persons living in certain geographic areas, institutionalised etc) were excluded. Once the population to be surveyed was defined, it had to

be represented in the physical form from which the sample of the required type could be selected. To draw the sample, the microdata of the 1989 population census were used as a sampling frame, given the aim of EFFS, the sampling unit was an individual. Despite relatively old, it was the only dataset ensuring the coverage of Estonian population. Besides providing a list of individuals, census data supplied each case in gross sample with a diverse background information [Katus, Puur 1993]. The availability of these characteristics is below used for the analysis of non-response and representativeness of the final sample, the presence of census information also excluded the incidence of fabricated interviews and other interviewer cheating.

The period that had elapsed since the date of the 1989 census, indeed involved some over- and undercoverage. On one hand, the sampling frame included persons who had died or emigrated from the country, on the other hand it did not include persons who had immigrated since 1989. Mostly, the overcoverage was excluded from the sample in the process of address verification. Despite it is difficult to give the exact estimate of undercoverage, given the general decrease in migration intensity over the past few years and the emphasis of the survey the micro-level analysis of demographic behaviour rather than on producing some aggregate estimates, the coverage problems can hardly be of significant relevance to aims of the survey.

### 3.2. Sampling Plan

The sample was to be representative of the eligible population, reflecting both its regional, ethnic, and social composition. As an aid in calculating the sampling fractions, the target population was broken into four categories: Estonians vs. non-Estonians, and persons ages 20-49 (born in 1944-1973) vs. ages 50-69 (born 1924-1943). The mentioned stratification scheme yielded four sample subpopulations or strata: Young Estonians, Old Estonians, Young Non-Estonians, and Old Non-Estonians. The factors that would affect the loss in cases between the sample and completed interviews were expected to be different across the strata. The primary task of stratification was to compensate these differences.

From the initial randomly drawn list of names from the 1989 census, the number of completed interviews in 1993-94 were expected to be determined by four sources of loss: (1) Death, (2) Emigration from Estonia, (3) Failure to Locate Current Address, (4) Non-Completion of Interview of those whose addresses were located. The losses due to the various factors were estimated as described below. In order to avoid the need to go back into the original census lists to choose more cases to achieve the target number of completed interviews, these estimates of losses were intentionally conservative.

(1) *Death*. Mortality estimates were based on life tables for Estonia for 1988-89 [Katus and Puur 1992], assuming a 5-year interval between the census date and the date of interview. They were calculated separately for women from the age group 15-45 to the group 20-49 and from the age group 45-65 to the group 50-69. The estimates were: Young Estonians and Non-Estonians, 0.5 per cent; Old Estonians and Non-Estonians, 5 per cent.

(2) *Emigration*. Based on estimates of recent emigration from Estonia [Sakkeus 1991], emigration of Young Estonians, Old Estonians, and Old Non-Estonians was assumed to be 1 per cent; and of Young Non-Estonians, 5 per cent.



(3) *Not Locating Address*. Based on the experience of *Saar Poll* and test cases by the Institute of Sociology, it was assumed that losses due to the failure to locate current addresses would be 15 per cent from each stratum.

(4) *Non-completion of Interviews*. Based on recent experience, it was estimated that 75 per cent of the Estonians and 65 per cent of the Non-Estonians who were located and contacted would complete the survey. Given the length of the planned field period in this survey, it was expected that the actual completion rates would be somewhat higher because of the opportunity to prolong efforts to contact respondents and to schedule interviews.

The calculation of losses from death, emigration and non-locations was additive and thus summed up to be: 17 per cent of Young Estonians, 21 per cent of Young Non-Estonians, 21 per cent of Old Estonians, and 21 per cent of Old Non-Estonians. The non-completion estimate was then multiplied by the proportion assumed likely to participate in the survey within each of the four groups after the first three effects had been applied. In order to obtain 5,000 completed interviews, with the respondents distributed by age and nationality as in the Estonian 1989 census, the following number of names was estimated to be obtained from the 1989 census files:

Table 1. STRATIFICATION OF THE EFFS SAMPLE

Sample Strata	Gross Sample Size	Estimated Net Sample Size	Sampling Ratio, %
Young Estonians	2,952	1,802	1.53
Old Estonians	2,006	1,182	1.56
Young Non-Estonians	2,570	1,315	1.87
Old Non-Estonians	1,370	701	1.74
All Strata	8,898	5,000	1.65

The uneven sampling ratios across strata established for purposes of making the sample more representative of the target population, not for purposes of developing a sample that was disproportionate to the population. Whatever the procedure, the selection of cases was decided to be without geographic stratification or clustering.

### 3.3. Stages of Sample Selection

The implementation of the sampling plan built on the availability of the individual-level data from the 1989 population census. An actual drawing of the sample occurred in two stages: first, a stratified random sample of 9,000 eligible persons was drawn from the census microdata, and an effort was made to obtain current addresses for all of these persons; second, cases were selected randomly in batches from the sample to be sent to *Saar Poll* for contacting and interviewing.

*First Stage of Sample Selection*. As the first step, using a simple random selection procedure within each stratum, 9,000 records were drawn from the census microdata. In round numbers, it was reasonable to draw 3,000 young Estonians, 2,000 Old Estonians, 2,600 Young Non-

Estonians and 1,400 Old Non-Estonians. The records were drawn from the universe of long-form lists, so that data from all variables in the 1989 census for individuals and households could be integrated with the survey results for purposes of analysis. This long-form was administered to a 25 per cent random sample of households in the census. Analysis showed that the characteristics of households in the long-form 25 per cent sample corresponded very closely with the characteristics of households in the short-form that was administered to the remaining households.

As the next step, the census records were delivered to the survey organisation. For the names and addresses of the individuals were not directly available from the datafiles, they were checked from original census lists in the archive at Statistics Estonia, current addresses for potential respondents were sought from the central address bureau. The link between datafiles and census lists was established by means of portfolio and list numbers. For 106 out of the initial 9,000 records, the list with addresses was missing from the folio in the archives. These 106 are considered losses from initial sample due to missing address information in census records.

For the remaining 8,894 persons, current addresses were sought from the central address bureau in Tallinn. Addresses were found in the address bureau for 7,676 individuals. The remaining 1,218 individuals fell into three categories (according to information in the address bureau records): 320 had emigrated, 136 had died, and for 762 there was no information (no address). The 456 individuals who had emigrated or had died were regarded as an overcoverage of the sampling frame and thus excluded. The 762 for whom no address was available in the central address bureau remained in the sample; however for cost considerations, they were given somewhat lower priority for selection in the second stage of the sample.

*Second Stage of Sample Selection.* The first stage of the sample selection and initial locating of respondents left 6,914 persons for whom a current address had been found at the central address bureau, and 762 for whom no current address had been identified. This did not mean that all persons with listed addresses would actually be found at those addresses, nor did it mean that all persons for whom no address was found in the central address bureau could not be found later by the interviewers at their census addresses. However, for the sake of cost, it was determined to select the first 5,000 cases to be interviewed from among those for whom current addresses had been identified. After approximately 3,000 interviews had been completed and a non-interview rate could be estimated for various categories of the population, additional cases were assigned to the interviewers. These cases were drawn at random from the remaining eligible persons in the sample, including those for whom no address had as yet been verified. First, 227 names were selected, then another 411, and finally 650 more.

In all, 6,212 eligible respondents were selected and sent to the survey organisation to be contacted and interviewed. Interviewers were instructed that if they could not locate the respondent at the listed address, then they should ask the local authorities (community administration, housing office etc) if they had more recent information. In the case of respondents for whom no address was found in the central address bureau, the interviewers were instructed to proceed on the basis of the address listed in the census (the "last known address"). If the interviewers found that the person in the sample had moved to another address, then an effort was made to find the individual at another address (if necessary, by assigning the case to a different interviewer).

### 3.4. Final Disposition of Cases

Table 2 shows the final disposition of cases that were assigned to the survey organisation. In all, 6,212 cases were assigned to the survey organisation. Of those, 5,021 interviews were completed and 1,194 were not completed. Of the latter category, in 50 cases no initiative was taken to contact the respondents.

Table 2. FINAL DISPOSITION OF CASES SENT TO SURVEY ORGANISATION

Disposition	Number of Cases	Percentage of Cases
Total in Second-Stage Sample	6212	100.0
Interviewed	5021 <sup>a</sup>	80.8
Not Interviewed	1192	19.2
of which:		
Contacted	1142	18.4
No contact initiative	50	0.8

<sup>a</sup> Includes one person who was erroneously interviewed: had same name as the person who was intended to be included in the sample.

The final sample size brings the sampling ratio 0.93 percent which means that each individual constituting a data unit can be regarded to represent 107 persons in the target population.

### 3.5. Reasons for Not Completing Interviews

Of the 1,142 persons who were contacted but not interviewed, the reasons for the noninterview are listed in Table 3. The most common reasons of non-response were refusals and failure to locate the respondent. For this purpose, we count as a "refusal" those cases that were "final" refusals. In some cases, the respondent's immediate response was clear and unequivocal: "Do not contact me again," "I have absolutely no interest in participating in this study," etc. These respondents, totalling 288 were not contacted again. In other cases, the initial refusal was considered "soft": "I am too busy," "I am not available now," "I am not sure I want to participate," "I am not feeling well now," etc.

Table 3. REASONS FOR NON-INTERVIEW OF THOSE WHO WERE CONTACTED

Reason for Non-Interview	Number of Cases	Percentage of Cases
Dead	43	3.7
Ill/Disabled	117 <sup>a</sup>	0.2
Emigrated	188 <sup>b</sup>	16.5
Not located	380 <sup>c</sup>	33.3
Refused interview	414	36.3
Total	1,142	100.0

<sup>a</sup> Includes 19 who were aggressively or incoherently alcoholic (established by interviewer on the basis of efforts to contact as well as actual contacts with persons), 61 who were seriously ill or disabled (in hospital, paralyzed, deaf or mute, had cancer, or mentally ill), and 34 who refused to participate on account of illness but the nature of the illness was not specified.

<sup>b</sup> Includes 72 who had emigrated from Estonia permanently and 116 who had emigrated temporarily (during entire field stage of survey).

<sup>c</sup> Includes two persons who were reported to be in prison. In general, includes persons for whom an effort was made to contact them (by post, telephone, or in person) but from whom no response or communication was received.

In such cases, the apparent uncertainty or temporary nature of the refusal was taken only as a soft refusal and additional efforts were made to secure an interview. The respondent was sent a refusal conversion letter, and whenever necessary, the case was reassigned to another interviewer. Many such efforts led to completed interviews, however 126 soft refusals remained unconverted.

It is noteworthy that losses from non-location were twice as high as the losses from emigration. The "did not locate" category resulted primarily from the inaccuracy of residence registration, particularly from the discrepancy between real and reported official addresses (based on so-called soviet *propiska*). The category includes all cases in which an effort was undertaken to contact the respondent by post, mail, or going to the place of last known address but for which no communication at all was received from the respondent. In many cases they were temporarily living on another address which the interviewer failed to establish, including temporary absence for summer season. Some respondents had never lived at the address indicated by the address bureau and/or recorded at the census. In some cases, the individuals may intentionally have avoided contact.

In our opinion this kind of a specificity should be taken into account in sampling plans for countries of the former Soviet Union.

### **3.6. Response Rates**

According to the outcome, the crude response rate for the EFFS was  $5021 : 6212 = .808$ , or 80.8 per cent. This assumes that all cases that were assigned to the survey organisation were "at risk" of being interviewed.

However, as enumerated above, some of the cases assigned to the survey organisation included persons who were later discovered to be deceased (43), were too ill to participate in the survey (117), had emigrated (188), or were not actually contacted (50). Regarding the non-completion caused by death, illness and emigration as overcoverage, we can recalculate the rate by including only those respondents who can be considered to be "at risk" of being interviewed and receive the corrected response rate  $5021 : 5815 = .863$ , or 86.3 per cent.

Finally, it is worthwhile to calculate the refusal rate: of those who were contacted (or to whom an effort to establish contact was addressed), what proportion directly declined to be interviewed? This calculation does not include persons who initially refused but who ultimately completed the interview. Nor does it include persons in the "did not locate" category. The refusal rate is  $414 : 6212 = .067$ , or 6.7 per cent.

Compared to other major surveys recently conducted in Estonia (Labour Force Survey, n=9608), the EFFS has slightly higher non-response rates. Higher response rates of the mentioned survey is explained by better performance of interviewer network (ELFS used the newly-established governmental network based on county statistical bureaux).

### 3.7. Representativeness of the Sample

The main concern with representativeness is not to have an exact replication of the target population but to identify any systematic ways in which the EFFS respondents may be different from the referent population. For this purpose two kinds of information are applied. First, the crude response rates for different categories of the eligible population are examined: what percentage of persons in the sample after the first stage completed the survey? The special interest is in whether the probability of completing the survey was associated with factors that were likely to be important in the data analysis. Specifically, the completion rate is examined in relation to region, urban-rural residence, ethnicity, education, age, and the number of children ever born. Next, direct comparisons of the characteristics of three "populations": the 1989 referent population of Estonia, the EFFS respondents, and the EFFS non-respondents are performed, applying the same set of variables. In this exercise, an advantage is taken of the data from the 1989 census which is equally available for all three.

*Crude Response Rates by Characteristics of Person.* Table 2 presents the response and non-response rates for some categories of referent population that are likely to be important in data analysis. If we examine the crude completion rates, we find little evidence of selection bias.

As expected at the time the sample was designed, the completion rate was higher for Estonians than for Non-Estonians. However, the difference was smaller than expected, and the overall completion rate for both groups was larger than anticipated by about ten percentage points. Most of the difference between these subgroups was due to greater emigration propensity among non-Estonians, and to a lesser extent higher prevalence of non-location. Estonians were found somewhat more frequently refusing to answer.

Rural residents were more likely to complete the survey than urban residents. Somewhat opposite to expectations, this is not explained by ethnic differences. Among non-Estonians, 77.0 per cent of urban residents completed the survey, compared to 75.6 per cent of rural residents. Among Estonians, on the other hand, 81.2 per cent of urban residents completed the survey, compared to 87.5 per cent of rural residents. Thus, the higher completion rates among rural residents were found only among Estonians. The main reason for this result was to an equal extent the higher refusal and non-location rate in urban areas, and to somewhat lesser extent the emigration. Among non-Estonians these factors were balanced by higher rates of emigration in rural areas.

The response rates by county are also suggestive. In Järva, Põlva, Rapla and Viljandi counties the rate was fairly above 90 per cent, the lowest rates were found in Harju (in which Tallinn is located) and Tartu counties.

Better educated respondents had slightly higher response rate, mostly due to lower prevalence of death and illness among more educated with other reasons being equal. As the decrease in emigration and non-location balanced the increase in death and illness, the response rate was almost equal across age groups. That the completion rate is the same for the older cohorts as for the younger ones is notable because of the evidence from interviewer reports that older respondents were more likely to have difficulties in the interview process.

By number of children, women with more children had higher completion rate, however, the differences in completion rate by number of children are small and not related to differences by cohort.

Table 4. RESPONSE RATES BY RESPONDENT CHARACTERISTICS

Respondent Characteristic in 1989 Census	Crude Response Rate, %	Crude Non-Response Rate, %	Corrected Non-Response Rate, %	Refusal Rate, %	Non-Location Rate, %
Whole Sample	80.8	19.2	12.8	6.7	6.1
Born 1969-1973	79.2	20.8	13.8	5.3	8.5
Born 1964-1968	79.8	20.2	13.5	5.3	8.2
Born 1959-1963	82.0	18.0	13.6	6.6	7.0
Born 1954-1958	80.0	20.0	14.6	6.9	7.7
Born 1949-1953	84.6	15.4	10.4	4.6	5.8
Born 1944-1948	77.8	22.2	15.3	8.4	6.9
Born 1939-1943	83.3	16.7	12.4	8.0	4.4
Born 1934-1938	81.0	19.0	13.5	8.0	5.5
Born 1929-1933	80.5	19.5	10.6	6.6	4.0
Born 1924-1928	79.9	20.1	10.2	6.7	3.5
Estonians	83.3	16.7	12.5	7.1	5.4
Non-Estonians	76.9	23.1	13.2	6.0	7.2
Urban	79.2	20.8	14.1	7.3	6.8
Rural	86.0	14.0	8.5	4.7	3.8
Harjumaa	76.4	23.6	16.5	8.3	8.2
Hiiumaa	80.5	19.5	17.1	19.8	7.3
Ida-Virumaa	81.5	18.5	9.5	4.6	4.9
Jõgevamaa	86.8	13.2	10.5	7.9	2.6
Järvamaa	92.0	8.0	4.9	4.3	0.6
Läänemaa	85.3	14.7	5.9	4.4	1.5
Lääne-Virumaa	85.4	14.6	11.1	5.0	6.1
Põlvamaa	91.6	8.4	4.2	2.5	1.7
Pärnumaa	83.3	16.7	12.0	6.5	5.5
Raplamaa	91.9	8.1	6.6	5.1	1.5
Saaremaa	84.4	15.6	7.8	4.3	3.5
Tartumaa	77.3	22.7	15.5	7.6	7.9
Valgamaa	82.8	17.2	11.2	6.7	4.5
Viljandimaa	91.8	8.2	6.1	3.9	2.2
Võrumaa	89.4	10.6	8.2	4.4	3.8
Higher Education	83.2	16.8	13.1	6.5	6.6
Secondary Educ	81.0	19.0	13.1	6.8	6.3
Less than Sec	79.6	20.4	12.1	6.6	5.5
Married	81.8	18.2	12.2	6.3	5.9
Never Married	77.7	22.3	14.3	7.4	6.9
Widowed	81.5	18.5	11.2	6.4	4.8
Divorced	80.4	19.6	14.1	7.4	6.7
Children Ever Born					
0	75.5	24.5	16.3	7.8	8.5
1	78.8	21.2	14.2	7.6	6.6
2	83.4	16.6	11.3	6.2	5.1
3	87.0	13.0	8.2	3.8	4.4
4 or more	87.3	12.7	7.4	4.9	2.5

The evidence on reasons of non-completion suggests that it may have resulted partly from lesser mobility of women with higher parity, and partly from the level of interest in the survey: women who had had children (and particularly women who had had several children) were more likely to take interest in a survey on family and fertility. The contacting letter sent to the respondents clearly identified the survey as a "family and fertility survey" and described the main topics to be covered (see the Appendix). Differential interviewer experience did not have any significant effect on response rate. Overall, observed differential rates are not likely to impart any notable bias to the data analysis.

*Comparisons of 1989 Eligible Population, FFS Respondents, And FFS Non-Respondents on Selected Characteristics.* Taking the advantage of the 1989 census, it is possible to compare the eligible population in 1989 with the EFFS respondents and non-respondents in order to evaluate the representativeness of the sample (Table 5).

By comparing the 1989 census data on total population with EFFS respondents, one observes that the characteristics of the EFFS respondents are very close to those of the eligible 1989 population. There seems little selectivity of the respondents relative to the population from which they were drawn. By comparing target population with non-respondents, one can see that those who were not interviewed are remarkable only for one characteristic: disproportionate numbers of the non-respondents were non-Estonians. Whereas 40 per cent of the eligible population were non-Estonians, 47 per cent of the non-respondents were non-Estonian. The result concerning the higher non-response among non-Estonians was anticipated in the sampling plan and led to an intentional over-sampling. As can be seen from the table, the distribution of EFFS respondents according to ethnicity approximates closely the distribution in the eligible population.

The figures for separate counties in column 3 also show a fair amount of variance from the column 1. This variance probably primarily reflects the small number of cases in the typical county and the different performance and efforts of individual interviewers rather than any real difference in the level of interest in the survey by respondents in the different regions.

On the whole, the final distribution of EFFS respondents approximates closely the eligible population. Further analysis of the material does not require any additional weights or adjustments to the EFFS results to make them more representative of the eligible population. Further analysis of the material does not require any additional weights or adjustments to make the results more representative of the target population.

Table 5. COMPARISON OF CHARACTERISTICS OF 1989 ELIGIBLE POPULATION, EFFS RESPONDENTS, AND EFFS NON-RESPONDENTS

Respondent Characteristic in 1989 Census	1989 Eligible Population, %	EFFS Respondents, % (n=5020)	EFFS Non-Respondents, % (n=1192)
Born 1969-73	9.8	9.2	10.2
Born 1964-68	9.4	9.9	10.5
Born 1959-63	11.0	10.1	9.3
Born 1954-58	11.4	9.9	10.4
Born 1949-53	10.9	9.6	7.4
Born 1944-48	9.1	9.6	11.6
Born 1939-43	9.4	10.9	9.1
Born 1934-38	9.8	10.6	10.5
Born 1929-33	9.7	10.0	10.2
Born 1924-28	9.6	10.1	10.7
Estonians	59.8	62.7	52.9
Non-Estonians	40.2	37.3	47.1
Urban	74.8	75.0	82.9
Rural	25.2	25.0	17.1
Harjumaa	40.0	39.0	50.8
Hiiumaa	0.7	0.7	0.7
Ida-Virumaa	14.9	15.1	14.5
Jõgevamaa	2.5	2.6	1.7
Järvamaa	2.6	3.0	1.1
Läänemaa	2.0	2.3	1.7
Lääne-Virumaa	4.8	4.4	3.2
Põlvamaa	2.1	2.2	0.8
Pärnumaa	6.2	6.4	5.4
Raplamaa	2.3	2.5	0.9
Saaremaa	2.3	2.4	1.8
Tartumaa	10.5	10.1	12.5
Valgamaa	2.5	2.2	1.9
Viljandimaa	3.9	4.2	1.6
Võrumaa	2.7	2.8	1.4
Higher Education	13.6	14.1	12.0
Secondary Education	54.2	53.8	53.4
Less than Sec. Educ.	32.2	32.0	34.6
Married	62.2	63.1	59.1
Never Married	19.9	18.9	22.9
Widowed	7.0	7.1	6.8
Divorced	10.8	10.8	11.1
0 Children Ever Born	24.2	22.6	30.8
1	27.1	26.6	30.0
2	34.9	35.0	29.3
3	9.7	10.9	6.9
4 or more	4.1	4.9	3.0



### 3.8. Linkage with Census Records

The use of the 1989 census microdata as a sampling frame made it possible to match the individual records from the two sources. The analysis performed on birthdate and other selected characteristics (marital status, number of live-born and surviving children, economic activity as of January 1989) revealed different levels of inconsistency ranging from 3.0 for year of birth to 11.8 percent of respondents for marital status and 13.6 per cent for usual activity. Being evidently caused by random errors at different stages of data collection, differences in birthdate did not display any systematic pattern. For the other characteristics with more extensive mismatch, some systematic bias was evident. As the survey addressed the issues with much greater detail than the census and because the comparison relates to the quite close past, there was no reason to ascribe the mismatch primarily to recall difficulties and lower accuracy of the survey. To provide an insight into the issue, the cross-classification of two sources is performed (Table 6).

Compared to the survey, the census over-estimates the number of singles and under-estimates the number of persons living in partnership. Cross-tabulation of census and survey responses revealed a substantial number of cohabiting respondents classified into single in the census. For other categories of marital status, the mismatch was smaller. Regarding economic activity, census appeared to over-estimate the activity rate by classifying women on maternity leave into economically active. Re-classification of the latter women revealed highly consistent figures from both sources. The mismatch in the number of children was primarily caused by the omission of children who had died in early infancy from the census count. As the described inconsistencies showed no unexplained variation across basic characteristics of the respondent, they should not be a major concern for data analysis. On the other hand, the mismatch between the two sources allows for the evaluation of census concepts and definitions.

Table 6. COMPARISON OF SELECTED CHARACTERISTICS OF RESPONDENTS FROM THE 1989 CENSUS AND EFFS SURVEY, AS OF JANUARY 1989

Category	Censusdata, %	Surveydata, %
Marital status		
Married	63.1	66.1
Never-married	18.9	15.1
Widowed	7.1	6.7
Divorced	10.9	12.1
Economic activity		
Active	79.9	71.7
Non-Active	21.1	28.3

## 4. DATA COLLECTION

Characteristic to the Soviet-type statistical system, out of three main ways of data collection - the census, vital statistics and surveys - the latter was obviously the most underdeveloped. Given the lack of experience and skills within Statistical Office newly established commercial survey agencies received certain competitive advantages. As the old network at Statistical

Office had been dissolved in the beginning of economic transition, the Working Group decided to use the interviewers from one these new agencies.

#### **4.1. Interviewer Training**

Taking into account the complexity of the survey and the lack of professional interviewers working on regular basis, the training was considered of key importance to the quality of data collection. In addition to the general training provided by *Saar Poll*, all interviewers had to go through a special 2-day training seminar on EFFS. To ensure the similar quality, all seminars were provided directly by EKDK researchers who were responsible for the development of survey instrument. The training seminars were held from December 1993-February 1994 in five rounds in Tallinn, Tartu and Narva, separately for Estonian and Russian-speaking interviewers.

The seminars concentrated on the concepts and definitions on which the questionnaire was designed, specific requirements concerning the collection of life history data (especially the importance of recalling the dates of the events and consistencies between different modules of the questionnaire) and procedural aspects. Apart from the usual practice of the survey organisation, at the end of the seminar each potential interviewer made a pretest on the questionnaire, after which she was selected for the task. The pretest consisted of a 30-60 minute interview where one of EKDK researchers took the role of the respondent with specially prepared life history including complex cases and various inconsistencies. The selection was based on general skills as well as the ability of interviewer to solve the cases and detect inconsistencies. Given the uncommonness of such a pretest for the survey organisation, no interviewers passed the test after the first seminar calling forth critique towards the individual testing; later the success rate ranged between 50-75 percent.

#### **4.2. Organisation of the Fieldwork**

The experience gained during pretests with different candidates for survey agency had indicated clearly the inadequacy of supervisory routines. As the local-level coordination turned to be weak, it was decided to operate the network directly from Tallinn without connecting links between the interviewer and the centre.

After the seminar the names and addresses of potential respondents were distributed to interviewers by their region and language. With a few exceptions, the interviewers were matched to the respondents according to language and nationality. That is, Estonians were generally interviewed by persons whose native language was Estonian, and Russians were interviewed by persons whose native language was Russian. As far as possible efforts were made to ensure that interviewer and respondent did not know each other, in one case the respondent and interviewer turned to be the same person.

After having received the names interviewers were to determine the contacting schedule for the entire field period. This procedure was to combine the autonomy of the individual interviewer with necessary central control under minimum subsequent exchange of information. According to agreed schedule, the survey organisation started to send out the contact letters together with

the address and phone number of their interviewer. The letter provided potential respondents with basic information about the survey and asked the respondent to designate the method of contacting (to send, on a separate response card, her phone number or suggested date of interview, and a new address, if changed) as well as a preference for the language at the interview. If the answer on language did not match the census information, the interviewer was changed to the appropriate one. If the non-correspondence of languages occurred during the direct contact by the interviewer, another interviewer was substituted. At the end of the letter, respondents were assured confidentiality of their responses.

Regarding the respondents, in total 2894 or 57.6 per cent send the response card back. There was virtually no difference between urban and rural dwellers while native-born respondents had return rate higher for 8 per cent. Older respondents and those with university degree were somewhat more likely to return the card. Despite relatively small variation, returning the card proved to be quite good indicator of the respondents attitude towards the survey and predictor of potential difficulties the interviewer encountered. Excessive return rate in January-February may reflect the timing of the media campaign to publicize the launching of this survey. To convey the information to the public knowledge TV and radio broadcasts were arranged, several articles appeared in the central and local newspapers about the survey to be carried out. Regarding the mode of contacting 41.3 per cent send their phone number, 19.7 percent called themselves to the interviewer and 35.4 per cent settled the time for the interviewer to come. In remaining 3.5 per cent of cases, the response card included respondent's new address. Thus beyond its informative function, the introductory letter seems to have contributed to the saving of interviewers time and resources.

In urban areas, interviewers were permitted to contact the respondents directly before receiving a contact letter back but only 10 days after the first contact letter was posted. In rural areas, in order to save transportation costs, the interviewer was advised to wait 20 days from the first letter before taking the next step; if they had not received any contact by that time, a second letter was sent out and after that the interviewer herself had to make the contact.

If the individual responded to the contact letter or otherwise agreed to be interviewed, then the interview was scheduled and subsequently completed (in most cases). The data were collected by face-to-face interviews using a paper-and-pencil method. Apart from normal procedure, there were three respondents in the survey who refused to make an interview but agreed to complete the questionnaire by themselves. In these cases the interviewer delivered the questionnaire, explained it to the respondent and arranged a time to return and pick up the completed form. Out of these three questionnaires, two later accepted and one rejected for incompleteness. In each successful contact the interviewer informed the survey organisation (mostly by phone); to monitor the progress of fieldwork the survey organisation updated weekly the sample file with the information on incoming questionnaires and results of contacting.

If information was received that the person had deceased or had emigrated, the effort to interview ended. The same was true when the individual was found to be institutionalized: in hospital or (in two cases) in prison. If the respondent stated that she was too ill to participate, no further effort was made to complete the interview. Persons who communicated with the survey organisation in writing, by telephone, or in person that they did not wish to be interviewed were treated as "refusals." Those for whom the initial refusal was "soft" were asked again whether they would agree to be interviewed. Some of these persons were ultimately interviewed.

In all cases the interviewers documented in detail their contacting efforts and the respective results (dates and time of visits and calls to the respondent, motives for refusal and conversion efforts, additional information received from relatives, neighbours, local authorities, housing office etc) on a special protocol. When the interviewer considered that all means to secure an interview were exhausted, the protocol was sent to survey organisation. In the survey organisation the protocols were screened, and whenever necessary returned to work with relevant instructions. The final decision to accept (or reject) the non-response was made by EFFS Working Group.

### 4.3. Number of Interviewers

As most of interviewers were working for a survey organisation on the part-time basis, the Working Group insisted on doing the work with smaller number of interviewers. As a compromise between quality requirements and funding available for transportation, the minimum number of interviews per interviewer was scheduled around 50 during the 5 months period. As a whole 70 interviewers were planned to be used initially.

Table 7. SELECTED CHARACTERISTICS OF INTERVIEW BY INTERVIEWER EXPERIENCE

Characteristics of Interview	Number of Interviews Conducted per Interviewer			
	<50	50-99	100+	Total
Duration of Interview, %				
Less than 30 min	0.2	0.6	0.0	0.4
30-59 min	7.6	9.5	21.7	11.0
60-89 min	38.2	41.5	59.3	43.6
90-119 min	33.0	34.5	16.2	31.3
120+ min	21.1	13.9	2.7	13.7
Average Duration, min	94.5	89.1	72.2	87.5
Interviewer Assessment of Respondent's Interest Towards the Survey, %				
Very Interested	19.8	24.5	22.7	23.2
Somewhat Interested	55.5	56.5	65.5	57.7
Somewhat Uninterested	22.1	17.3	10.4	17.2
Very Uninterested	2.7	1.8	1.4	1.9
Crude Non-Response Rate, %	15.5	20.6	13.6	18.5
Prevalence of Date-Specific Non-Response, %				
Year	6.1	5.2	2.8	4.9
Month	22.4	18.7	12.8	18.9

In practice, however, 93 female interviewers were used out of whom 65 worked primarily in Estonian and 28 in Russian language. Out of these 93, 36 did less than 50 interviews, 39 did between 50 and 75 interviews and 15 more than 75 interviews. Three interviewers quitted in the middle of the fieldwork, having completed fewer than 10 interviews, the respondents assigned to these persons were distributed among others. The maximum number of contacted persons for an interviewer was 294, of which the interviewer completed 234 interviews.

Data presented in Table 7 suggest a clear relationship between the experience and interviewer performance: the more interviews, the better results. Besides that it is interesting to note that better results were achieved by newly-recruited interviewers rather than by those having

worked for *Saar Poll* before, and being evidently used to the requirements of public opinion and other sociological surveys.

#### 4.4. Month of Interview

Originally the field period was scheduled to last about five months, from January through May 1994. However for a variety of reasons, it was not possible to complete the interviewing by that deadline. The main part of the fieldwork was completed in January through August 1994, inclusive, with a few cases being completed in September through November. Given the nature of this survey and the funds that were available, it was not considered essential to concentrate the interviews into a short span of time.

Table 8. DISTRIBUTION OF INTERVIEWS BY MONTH OF COMPLETION

Month	Number of Interviews	Cumulative Percentage
January	128	2.5
February	837	19.2
March	1044	40.0
April	880	57.5
May	861	74.7
June	566	86.0
July	296	91.8
August	264	97.1
September	93	99.0
October	32	99.6
November	20	100.0

#### 4.5. Location of Interview

Nearly three-fourths (73.5 per cent) of the interviews were completed at the respondent's home, with the remainder occurring at the respondent's place of work or school (10.1 per cent), at the survey organisation (0.3 per cent), or elsewhere (16.0 per cent). These "other" places most often included the interviewers home or working place. Variation in the location interview across categories of respondents is small with older persons being slightly more frequently interviewed at home and younger and better educated ones at workplaces.

In 88.2 per cent of cases no other persons were present at the interview, in 2.4 per cent of cases the interviewer reported the presence of the spouse, in 8.2 per cent the presence of children and in 2.4 per cent of cases the presence of other relatives or non-relatives.

#### 4.6. Length of Interview

The life history format of the survey and the amount of information to be recorded resulted in relatively long interviews. The average number of life events recorded per respondent (operationalised here as the total number of columns filled in life history tables) was around

17-18, with practically no variation starting from the 1959-1963 cohort. The biggest average number of events was recorded in migration section and the smallest in the partnership section.

In the process of questionnaire development, the duration of an interview was expected to be within the limits of an hour and a half. In the survey, the average actual length of an interview was 87.5 minutes. Three fourths of the interviews lasted between one and two hours with hour and a half being the most common duration, in 30 cases the interview was interrupted and continued later.

Systematic variation in the length of the interview was observed according to categories of respondent. Differences by age ranged from an average of 75.7 minutes for persons born in 1969-1973, to 94.2 minutes for persons born in 1924-28. The percentage of interviews with duration over two hours among the oldest respondents was more than three times higher than among the youngest, the age-pattern of very short interviews reveals the opposite. On the average, for respondents with secondary and higher it took nearly ten minutes less to complete an interview than for their less educated counterparts. The interviews taken in the beginning of field period (January-February) lasted about 100 minutes while interviews made towards the end field period June-September lasted about 80 minutes. As discussed above, the length is strongly related to the interviewer experience. On one hand, for interviewers who made less than 50 interviews, the average duration of interview was 94.5 minutes. On the other hand for those who interviewed more than hundred respondents, it took 72.2 minutes to complete a questionnaire. Among the first category, the percentage of interviews which lasted more than one and a half hours was 54.1 percent while for the most experienced interviewers the percentage of thus interviews was only 18.9 percent.

Before conducting the fieldwork a concern about the length of the survey was expressed by the foreign advisors. Comparison of interview characteristics by duration presented in Table 10 does not reveal the longer duration having involved any significant deterioration in quality. The positive relationship between interviewer assessment and length reflects the longer time being associated with recall difficulties rather than with other characteristics.

Table 9. SELECTED CHARACTERISTICS OF INTERVIEW BY DURATION OF INTERVIEW

Characteristics of Interview	Duration of Interview, Minutes				Total
	-59	60-89	90-119	120+	
Interviewer Assessment of How Well the Interview Proceeded, %					
Very Well	56.6	41.4	35.9	39.2	41.1
Fairly Well	36.4	47.4	48.4	38.3	45.5
Some Difficulties	6.1	9.7	12.2	15.6	10.9
Big Difficulties	0.9	1.5	3.5	6.9	2.5
Proportion of Interviews that Required Second Contact, %					
	9.9	11.2	12.2	12.7	11.6
Index of Mismatch With Individual Census Characteristics <sup>a</sup>					
	113.1	99.7	100.4	95.7	100.0

<sup>a</sup> Index is based on summary prevalence of mismatch between the characteristics of individual respondents in the 1989 census and survey; characteristics used to construct the index include date of birth, marital status, number of live-born and surviving children, educational attainment and economic activity as of January 1989.

Apart from objective characteristics the low interest of the respondent towards the survey, and especially difficulties reported by the interviewer are accompanied with substantial increase in

the duration of the interview. Given the small difference in the number of life events in the latter interviews the difficulties seems to result personal characteristics of the respondent rather than from the excessive amount of detail to be recorded and respectively lengthy interviews.

#### **4.7. Quality of Interviews: Interviewer Assessments**

At the end of each interview, the interviewers were required to answer a few questions about the respondent, such as how interested the respondent was, and how well the interview seemed to proceed. While there is not much evidence in the scientific literature about the relation between interviewer perceptions of the quality of the interview and the actual quality of the interview (in terms of the accuracy of the results), it is of interest to obtain some idea of the quality of the social interaction that took place.

On average, about 23.2 per cent of the respondents were judged to be very interested in the survey, 57.7 per cent somewhat interested, 17.2 per cent somewhat uninterested, and only 1.9 per cent very uninterested. There was some variance in the perceived level of interest in the survey by respondent characteristics. The percentage of very interested was higher among native and better educated respondents. Older respondents were judged to be only slightly less interested than younger respondents. For example, among respondents born in 1924-28, 22 per cent were said to be very interested and 55 per cent to be somewhat interested, while for the respondents born in 1964-68, 21 per cent were said to be very interested and 61 per cent to be somewhat interested. Interestingly, the number of children was not associated with respondents attitude towards the survey.

The interviewers' evaluation of how well the interview proceeded follows a somewhat different pattern. On average, about 41.1 per cent of the interviews were said to have gone very well, and 45.5 per cent well, compared to 10.9 per cent that reportedly proceeded with some difficulties and 2.5 per cent with great difficulties. However, the interviews that proceeded with difficulties were concentrated among the older respondents, especially those born before 1944. Whereas only 41 per cent of the respondents were born before 1944, 71 per cent of the interviews that proceeded with great difficulty were among respondents born before 1944. Nevertheless, overall the interviews were judged to have proceeded quite well, and the greater difficulties for older respondents are likely to have been caused mainly by the difficulty of recalling detailed dates of events. Even for the oldest respondents (the cohort born 1924-28), 78 per cent of the interviews were judged to have proceeded very well or well, compared to 17 per cent that proceeded with some difficulties and 4.9 per cent with great difficulties. The variation by educational attainment follows the pattern similar to that observed in the case of attitude.

An analysis of the relationship between interviewer experience on this survey and the quality of the interviews reveals that the more interviews any individual conducted, the greater the reported interest of the respondents in the survey and the fewer the reported difficulties in carrying it out.

## 5. CODING, DATA ENTRY AND EDITING

To ensure the quality, the coding and data entry was performed by a separate unit under the supervision of EKDK. Before entering the data, codes were inserted in the questionnaires for occupation, educational qualification, ethnicity, language and regions. Before entering the data the codes were inserted in the forms for occupation, educational qualification, ethnicity, language and regions. Occupations were coded using the second level of ISCO-88, codes for educational qualifications were based on ISCED. The coding was complicated by the fact that Statistics Estonia has not completed the transfer to international classifications and that respective manuals and coding schemes, not to mention the computer-assisted coding facilities, were not available. Given the situation, in the framework of EFFS necessary instruction materials had to be prepared.

The procedure did not foresee manual pre-editing of the questionnaires. Data were entered by means of the *ISSA*-based data entry program which utilised took care of checks for the principal inconsistencies between the recorded answers. If inconsistencies were encountered, they were either resolved consulting related answers in the questionnaire or the case was sent back to the interviewer for clarification. In latter cases the interviewer received copies of these pages including inconsistencies or omissions and was asked to re-contact the respondent. Costs for the second contact were not compensated to the interviewer and the piece-rate was reduced by 20 per cent.

At the beginning of the fieldwork, first two completed interviews of every interviewer were subject to thorough reviewal. Based on the analysis of typical errors discovered in these questionnaires, the special instructions were developed to send to all interviewers. Despite efforts taken to reduce the error rate, almost 11.6 per cent (582) of the completed interviews needed a second contact, analysis did not reveal any systematic variation return rate by the characteristics of the respondent or interview situation. In the process of data entry 5 questionnaires were rejected due to incompleteness, four of these were done in institutions with mentally handicapped respondents and one was self-administered.

Once all the data were entered, the file was converted into *SPSS-Windows* format and systematically checked. Altogether the check programs covered several hundreds potential errors including those already incorporated in the data entry facility. Whenever the inconsistencies were detected, the file record was compared to questionnaire. In the overwhelming majority of cases, it proved possible to correct the problem by combining the information from different questions. Most of the errors had resulted from the interviewers inattentiveness, data entry errors accounted for about one tenth of edits. Regarding open-ended questions, the questionnaire design attempted generally to minimize the use "other - specify" approach. At the editing stage, selected part of the open-ended was checked and partially recoded. The case-by-case comparison of original and cleaned file revealed more than 20 thousand unit changes, altogether in the editing process more than three fourths of questionnaires were consulted. The analysis revealed systematic variation in the number of edits neither by respondent nor by interview characteristics.



## 5.1. Item-Specific Non-Response

The analysis reveals that the item-specific non-response varied for each question, although in most cases there was none of or it was infrequent. Where questions were applied to all respondents, the rate of item-specific non-response is identical to the number of missing data items, otherwise the rate should consider the number of responding persons to whom the question was addressed. Missing data resulted from the cases when the respondent chose not to answer a question or was unable to recall the information. Compared to recall difficulties item-specific refusals were encountered only in exceptional cases. It is noteworthy that regarding even the most intimate issues, there were only three persons who refused to provide information about the number of sexual intercourses during the last four weeks preceding to the interview. Out of these three, one interrupted the interview and was dropped from the count. The results on pregnancy history module were consistent with special abortion pretest which had helped to demonstrate that the population of Estonia was much more likely to respond accurately about some aspects of their reproductive behavior than is generally true in fertility surveys [Anderson, Katus, Puur, Silver 1993]. The stable ratio (2 to 1) between the number of pregnancies and live births across the cohorts suggests that there should not be major omissions on the topic.

From the viewpoint of life-history analysis, the quality of date-specific information is the most critical. Table 10 reveals that more than 95 per cent of the respondents were able to recall the year for all events asked. In major biographical modules year-specific non-response ranged from 0.0 to 0.4 per cent.

As to the month of event, the levels of non-response were expectedly higher. In general, 18.9 per cent of the respondents were unable to recall the exact month of event at some point in the interview. Frequently however, she could remember the part of the year the event took place, allowing the season codes to be applied. Regarding major biographical sections the non-response ranges between 0.5 and 0.9 per cent, somewhat higher prevalence of month-specific non-response (2.1 per cent) has also been recorded for other pregnancies. Nine tenths of the recall difficulties were located in parental home module, mostly due to inability of respondents to recall the birthdates for siblings who had died in early infancy.

Decomposition of date-specific non-response shows some systematic variation by characteristics of the respondent as well as by the dimensions of interview situation. Differences of 2-3 times can be observed across age, educational, native/non-native and parity dimensions with age-differences being discussed in detail below.

Compared to the average, interviewees, conducting more interviews, have evidently been more supportive in achieving twice lower date-specific non-response rates. Considering the modules in which the non-response is located, besides selective weakness of memory the observed patterns may also reflect the difference in the number of events to be recalled. Given the concentration of the overwhelming majority of the date-specific non-response in a few questions in the module of parental home which is less tightly related to the respondent's own biography. Due to the low level of non-response, its variation across respondent characteristics may be regarded irrelevant.

Table 10. DATE-SPECIFIC NON-RESPONSE BY RESPONDENT CHARACTERISTICS,  
MAIN BIOGRAPHICAL MODULES

Module/Respondent Characteristic	Year not Specified, %	Month not Specified, % <sup>a</sup>
Whole sample/All Modules	4.9	19.9
Partnership History	0.2	0.8
Childbearing History	0.1	0.5
Pregnancy History	0.4	2.1
Parental Home	4.4	17.1
Migration History	0.0	0.9
Living Arrangements History	0.1	0.8
Education and Work History	0.1	0.6
Estonians	3.9	16.4
Non-Estonians	7.6	26.1
Urban	5.5	20.6
Rural	4.5	18.4
Harjumaa	5.6	18.8
Hiiumaa	0.0	0.0
Ida-Virumaa	8.2	32.6
Jõgevamaa	7.1	29.9
Järvamaa	1.3	4.0
Läänemaa	15.8	41.2
Lääne-Virumaa	3.2	18.9
Põlvamaa	9.0	19.0
Pärnumaa	6.3	25.4
Raplamaa	4.6	8.1
Saaremaa	1.7	18.5
Tartumaa	1.4	1.1
Valgamaa	1.9	5.8
Viljandimaa	3.1	23.4
Võrumaa	0.0	4.4
Higher Education	1.4	14.3
Secondary Education	3.9	16.0
Less than Sec. Educ.	10.5	32.3
Married	2.4	10.5
Never Married	4.8	18.8
Widowed	9.0	29.9
Divorced	6.8	24.8
0 Children Ever Born	2.3	11.6
1	5.7	19.4
2	5.6	22.1
3	5.4	21.4
4 or more	7.3	25.7

<sup>a</sup> Does not include cases where respondent could not specify the exact month but was able to specify the part of the year event occurred.

## 5.2. Quality of Interviews for Older Cohorts

Extending of the upper age-limit of survey population from 50 to 69 indeed involved a risk of major trade-offs in data quality. For older respondents many events addressed by questions had taken place more than half a century ago, and therefore, it was natural to expect more problems in their case. The critical question which the survey had to answer was, whether these problems could be kept within acceptable limits. Data presented in Table 8 suggest that it has basically succeeded.

Table 11. SELECTED CHARACTERISTICS OF INTERVIEW BY BIRTH COHORT OF RESPONDENT

Characteristics of Interview	Birth Cohort										
	1924 1928	1929 1933	1934 1938	1939 1943	1944 1948	1949 1953	1954 1958	1959 1963	1964 1968	1969 1973	To- tal
Response Rate, %	79.9	80.5	81.0	83.3	77.8	84.6	80.0	82.0	79.8	79.2	80.8
Non-Response Rate, %	10.2	10.6	13.5	12.4	15.3	10.4	14.6	13.6	13.5	13.8	12.8
Interviewer Assessment of Respondent's Interest Towards the Survey, %											
Very Interested	22.4	21.9	21.7	22.4	22.8	23.2	27.6	24.3	21.1	24.8	23.2
Somewhat Interested	55.2	54.6	57.3	58.4	60.2	57.0	54.1	60.6	60.7	59.6	57.7
Somewhat Uninterested	20.0	20.3	17.4	17.9	15.5	19.2	16.1	14.1	16.4	14.3	17.2
Very Uninterested	2.4	3.2	3.6	1.3	1.5	0.6	2.2	1.0	1.8	1.3	1.9
Prevalence of Year-Specific Non-Response, %											
All Modules	11.8	7.9	7.4	6.2	4.8	4.5	3.5	3.4	0.9	0.1	4.9
Partnerships	0.2	0.0	0.4	0.2	0.2	0.0	0.0	0.2	0.4	0.0	0.2
Childbearing	0.2	0.2	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Pregnancies	1.2	0.2	0.6	0.6	0.2	0.4	0.2	0.2	0.0	0.0	0.4
Parental Home	10.1	7.7	6.6	5.2	4.4	4.1	3.2	2.8	0.4	0.1	4.4
Migration	0.0	0.2	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Living Arrangements	0.4	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Education and Work	0.2	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.0	0.2	0.1
Prevalence of Month-Specific Non-Response, %											
All Modules	33.9	31.3	30.8	26.0	19.9	18.1	14.7	10.3	6.3	6.0	18.9
Partnerships	1.7	0.9	1.6	1.2	0.7	1.1	0.4	0.7	0.4	0.0	0.8
Childbearing	1.5	1.5	0.8	0.8	0.0	0.2	0.2	0.2	0.0	0.0	0.5
Pregnancies	2.2	3.7	3.8	3.6	3.1	1.8	2.4	1.5	1.3	0.2	2.1
Parental Home	31.0	29.0	28.0	23.9	18.2	15.3	12.7	9.5	5.7	6.0	17.1
Migration	1.0	2.9	1.7	1.2	1.4	0.7	1.0	0.0	0.0	0.2	0.9
Living Arrangements	2.7	1.9	1.2	1.2	0.0	0.2	0.2	0.6	0.2	0.0	0.8
Education and Work	1.9	0.9	0.6	0.4	0.7	0.7	0.4	0.2	0.2	0.2	0.6
Proportion of Interviews that Required Second Contact, %											
	12.8	12.6	12.2	10.8	11.5	10.6	10.3	13.5	11.8	10.4	11.6
Index of Mismatch With Individual Census Characteristics, %											
	106.4	100.1	84.1	80.6	70.8	74.0	85.8	93.4	131.9	178.4	100.0

As noted above, due to different factors of non-response outbalancing each other, the completion rate remained practically stable across age-groups with somewhat lower levels among the oldest cohorts. For family and fertility matters are less close to daily concerns of older women, their interest towards the survey appeared to be slightly lower than among younger. Regarding the consistency of life histories, there were only slightly more interviews that required second contact among the older cohorts, however when considering the bigger number of life events, this pattern disappears.

The same is suggested by the number of edits introduced during data cleaning. The mismatch index between census and survey characteristics reveals rather the difference of definitions in census and the survey (particularly for marital status and economic activity) than the effect of age differences. Somewhat elderly-specific was only the mismatch in the number of children ever born but it is the inaccuracy problem in the census rather than in the survey data.

The only elderly-specific problem remained the date-specific non-response, but as it mostly cumulates into the module of parental home, it should not be a major concern for data analysis. Recall difficulties in that particular module seem to stem not as much from the time-span elapsed from the events but from the fact, that the events recorded in this module did not concern so much the respondent herself than her parents, particularly intergenerational differences in fertility levels and infant mortality.

### **5.3. Data File**

The data from the survey are available in *the* format *SPSS-Windows* system file, incorporating all variable and value labels. Into the file can also be incorporated the data from the 1989 census for each respondent.

At the editing stage, considering the low prevalence of unspecified dates and the need to facilitate later data processing, it was decided to impute the values for these dates for the basic life events (partnerships, pregnancies, migrations, work interruptions). No imputations were applied to dates recorded in parental home and living arrangements sections where the prevalence of unspecified dates was higher). Regarding the month of event, season codes were replaced by values representing the central month for the respective season. In case the month was fully unspecified, the *SPSS-Windows* Uniform-procedure was used to impute the missing value. Unspecified years of events were imputed manually consulting the timing of related events in other sections of the questionnaire. Doing so, it was in most cases possible to determine the period event had occurred with the precision of a few years. Due to low prevalence of missing values, these imputations did not have any significant effect on timing, sequencing and seasonality patterns of events concerned.

## REFERENCES

- Anderson, Barbara, Kalev Katus and Brian Silver (1994). Developments and Prospects for Population Statistics in Countries of the Former Soviet Union. *Population Index*, Vol.60, No.1, pp.4-20.
- Anderson, Barbara, Kalev Katus, Allan Puur and Brian Silver (1993). Characteristics of Women Having Abortions in Estonia. *Proceedings of the 22nd International Population Conference*. Montreal-Liège, IUSSP, Vol.1, pp.215-234.
- Anderson, Barbara, Kalev Katus, Allan Puur and Brian Silver (1994). The Validity of Survey Responses on Abortion: Evidence from Estonia. *Demography*, Vol.31, No.1, pp.115-132.
- Blossfeld, Hans-Peter, Alfred Hamerle and Karl Ulrich Mayer (1989). *Event History Analysis. Statistical Theory and Application in the Social Sciences*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, Publishers.
- Darsky, Leonid (1986). *Evolucija rozhdæemosti v SSSR po dannym obschesojuznyh vyborochnyh obsledovanii*. Paper presented to Deuxieme Seminaire Franco-Sovietique de Demographie, September 15-19. Suzdal.
- Katus, Kalev (1990). Demographic trends in Estonia throughout the centuries. *Yearbook of Population Research in Finland*. Helsinki, pp.50-66.
- Katus, Kalev (1991). *Fertility Trend and Regional Differences in Estonia*. RU, Series B, No.12. Tallinn, EKDK.
- Katus, Kalev (1994). Fertility transition in Estonia, Latvia and Lithuania. *Wolfgang Lutz, Sergei Scherbov and Andrei Volkov (Eds). Demographic Trends and Patterns in the Soviet Union Before 1991*. London and New York, IIASA, Routledge, pp.89-111.
- Katus, Kalev and Allan Puur (1992). *Eesti rahvastiku suremustrend elutabelite analüüsi põhjal*. RU, Series B, No.14. Tallinn, EKDK.
- Katus, Kalev and Allan Puur (1993). *The 1989 Estonian Population and Housing Census: Data Description*. RU, Series A, No.33. Tallinn, EKDK.
- Katus, Kalev, Allan Puur, Luule Sakkeus and Brian Silver (1995). *Eesti Pere- ja Sündimusuuring. Metodoloogiaülevaade*. RU, Series A, No.39. Tallinn, EKDK.
- Mitmekeelne Demograafiasõnastik. Eesti väljaanne*. (1993). Tallinn, EDA.
- Pavelson, Marje and Dmitri Mihhailov (1982). Sredovaya aktivnost' gorodskikh semei: na puti k territorial'noi obschsnosti. *Mikk Titma (Red). Sotsiologitskeskie issledovanija v Sovetskoi Pribaltike*. Vilnius, Institut Filosofii, Sotsiologii i Pravo AN LSSR, s. 264-280.

- Sakkeus, Luule (1991). *Post-War Migration Trends in Estonia*. RU, Series B, No.15. Tallinn, EKDK.
- Sakkeus, Luule (1994). The Baltic States. *Solon Ardittis (Ed). The Politics of East-West Migration*. New York, St. Martin's Press, pp.68-85.
- Titma, Mikk (1993). *Zhiznennyye puti odnogo pokolenia*. Moskva, Nauka.
- Tuma, Nancy, Michael Hannan (1984). *Social Dynamics: Models and Methods*. New York, Academic Press.
- UN ECE (1992a). *Fertility and Family Studies in the Countries of the ECE Region*. A Project Undertaken by the Population Activities Unit of the Economic Commission for Europe with Financial Support from the United Nations Population Fund. Geneva, United Nations.
- UN ECE (1992b). *Fertility and Family Studies in the Countries of the ECE Region*. Questionnaire and Codebook. New York, United Nations.
- UN ECE (1993). *Fertility and Family Studies in the Countries of the ECE Region*. Standard Recode Files and Standard Country Reports. Geneva, United Nations.

L I S A D  
*APPENDICES*

# EPSU KÜSITLUSINSTRUMENTAARIUM

## Eesti versioon

1. Küsimustik \*
2. Küsitlejajuhend \*
3. Protseduurijuhend
4. Kaardivihk
5. Kontaktkiri
6. Vastuskaart
7. Küsitlusprotokoll
8. Tänukaart

---

\* Ümbertrükk käesolevas väljaandes



# **EFFS SURVEY INSTRUMENT**

## **English version**

1. Questionnaire \*
2. Interviewer's manual \*
3. Fieldwork instructions
4. Response cards
5. Contact letter
6. Return card
7. Interview/refusal protocol
8. Thank-you card

---

\* Reprinted in the present volume

