

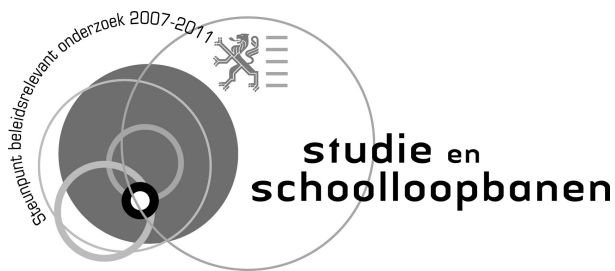


## Mag het iets meer zijn?

*Een toets van de effecten van een bevraging meer of minder door middel van een vergelijking van twee SONAR-cohorten*

Ignace Glorieux & Ilse Laurijssen

Onderzoek gefinancierd door de Vlaamse Regering in het kader van het programma 'Steunpunten voor Beleidsrelevant Onderzoek'



## **Mag het iets meer zijn?**

*Een toets van de effecten van een bevraging meer of minder door middel van een vergelijking van twee SONAR-cohorten*

Ignace Glorieux & Ilse Laurijssen

Promotor(en): Prof. Dr. Ignace Glorieux

Onderzoek in opdracht van de Vlaamse minister van Onderwijs en Vorming, in het kader van het programma 'Steunpunten voor Beleidsrelevant Onderzoek'

SSL-rapport nr. SSL/OD2/2011.33  
december 2011

**Voor meer informatie over deze publicatie:**

Steunpunt SSL, onderzoeksdomein 'Sociale Ongelijkheid in Schoolloopbanen, de Overgang van school naar werk en de Eerste jaren op de arbeidsmarkt (SOSOE)'

Auteurs: Ignace Glorieux & Ilse Laurijssen

adres: Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel

tel.: 02/6148150

fax: 02/6148140

e-mail: [torinfo@vub.ac.be](mailto:torinfo@vub.ac.be)

website: <http://www.steunpuntloopbanen.be>

# Mag het iets meer zijn? Een toets van de effecten van een bevraging meer of minder door middel van een vergelijking van twee SONAR-cohorten

Glorieux, I. & I. Laurijssen (2011):. Leuven: Steunpunt Studie en Schoolloopbanen, SSL-rapport nr SSL/OD2/2011.33.

## Samenvatting

De SONAR-enquête bij de geboortecohorte 1980 vond bij de respondenten na een initiële bevraging op 23 jaar meteen op 29 jaar plaats zonder dat een (tussenliggende) bevraging op 26 jaar plaatsvond, zoals dat wel het geval was voor de SONAR-cohorte van 1976. Dit impliceert een reconstructie van een langere periode van de loopbanen, aansluitend op de gegevens verworven op de leeftijd van 23. In dit rapport toetsen we (enkele van) de implicaties van een bevraging waarin een groter deel van de loopbaan aan bod moet komen. We gingen na of deze nieuwe dataverzameling dezelfde gedetailleerde informatie levert, en of de loopbanen met dezelfde accuraatheid kunnen worden gereconstrueerd. Als vergelijkingspunt nemen we de data van cohorte 1976 die eveneens bevraagd werd op 23 en 29 jaar, maar ook op 26 jaar. Daarbij vergeleken we de respons, het aantal geregistreerde periodes, de hoeveelheid “missing data”, maar ook bijvoorbeeld of bepaalde gegevens, en in het bijzonder de geregistreerde tijdsperiodes, meer “afgerond” worden.

## 1 Respons en evaluatie van het interview

De Vlaamse interuniversitaire onderzoeksgroep SONAR (afkorting voor Studiegroep van ONderwijs naar ARbeidsmarkt) startte in 1999 met het verzamelen van gegevens over het verloop van de transitie van onderwijs naar werk door middel van survey-onderzoek bij verschillende cohorten en op verschillende tijdstippen. Voor de Vlaamse jongeren geboren in 1976, 1978 en 1980 werden uit het Rijksregister aselechte steekproeven getrokken van ongeveer 3.000 personen per geboortjaar (en reserve-adressen). Al deze proefpersonen werden individueel bezocht en bevraagd op 23-jarige leeftijd (N=9.010). De respondenten geboren in 1976 en 1980 werden vervolgens opnieuw bevraagd op 26-jarige leeftijd in respectievelijk 2001 en 2004 (N=4.162). In 2005 vond nog een herbevraging op 29-jarige

leeftijd plaats van de cohorte geboren in 1976 (N=1.657). In 2009 werden de respondenten geboren in 1980 op 29-jarige leeftijd opnieuw bevestigd (N=1.922).

TABEL 1: SAMENSTELLING VAN DE SONAR-DATA

	<b>23-jarigen</b>	<b>26-jarigen</b>	<b>29-jarigen</b>
Geboren in 1976	C76(23) 1999 N=3015	C76(26) 2002 N=2060	C76(29) 2005 N=1657
Geboren in 1978	C78(23) 2001 N=3002	C78(26) 2004 N=2102	-
Geboren in 1980	C80(23) 2003 N=2993	-	c80(29) 2009 N=1922

In bovenstaande tabel staan al deze enquêtes opgelijst. In dit rapport maken we gebruik van de gegevens van de eerste en de laatste geboortecohorte. De respondenten uit die beide cohorten werden bevestigd op 29-jarige leeftijd. Voor de eerste geboortecohorte was dit een derde bevestiging, voor de laatste cohorte was dit een tweede bevestiging. In dit rapport staat centraal of het ontbreken van de tussenliggende bijkomende bevestiging (op 26 jaar) een verschil maakt, en wat de eventuele impact hiervan is op de kwaliteit van de gegevens. We zullen aan de hand van drie invalshoeken trachten hierop een zicht te krijgen. Ten eerste gaan we na hoe de uiteindelijke respons verschilt, en bekijken we of dat ook betekent dat de uiteindelijk bevestigde (overblijvende) groep op 29 jaar meer of minder selectief is op bepaalde achtergrondkenmerken.

Ten tweede bekijken we hoe respondenten zelf het interview ervaren, en hoe de enquêteurs de door hen bezochte en ondervraagde respondenten beoordeelden. Dit kan ons meer vertellen over (mogelijke verschillen in) de motivatie van de respondenten die nog meedoen aan de laatste bevestiging. We kunnen immers verwachten dat een extra bevestiging zich mogelijk vertaalt in een zekere interviewmoeheid, en dus verwachten we een hogere motivatie bij de respondenten van de laatste geboortecohorte. Anderzijds kan ook wel een mogelijke demotivatie verwacht worden bij de respondenten van de laatste geboortecohorte, in de mate dat hun interview langer duurt, en moeilijker is, omdat die respondenten voor de opdracht gesteld worden om een langer deel van hun loopbaan te reconstrueren (tussen 23 en 29 jaar).

En tot slot gaan we in op de vraag of met een bevestiging minder ook geen belangrijke informatie van de loopbaan verloren ging. De loopbanen worden in het SONAR-onderzoek uitvoerig in kaart gebracht door middel van kalenderregistraties (op maandelijkse basis). De schoolloopbaan wordt (hoofdzakelijk) retrospectief bevestigd; de arbeidsloopbaan werd bij elke enquête telkens verder aangevuld. In het bijzonder vergt de reconstructie van een langer deel van de loopbaan mogelijk niet alleen meer moeite en inzet (wat zich kan vertalen

in de motivatie, zoals hierboven vermeld), maar mogelijk is de kwaliteit van de bekomen informatie ook beperkter, doordat geheugeneffecten een rol spelen. We verwachten dan ook dat registraties van de loopbaan minder gedetailleerd worden opgesplitst, dat er meer ontbrekende informatie is, of ook dat de relevante periodes meer worden afgerond. Dit laatste gaan we in het laatste luik van dit rapport na.

## 1.1 Responscijfers en selectiviteit van de respons

In tabel 1 blijkt al een fundamenteel verschil in het uiteindelijk aantal respondenten bij de bevraging op 29-jarige leeftijd. In 2005 namen 1.657 respondenten van de eerste geboortecohorte deel aan het onderzoek op 29-jarige leeftijd. In 2009 werden 1.922 respondenten van de laatste cohorte bereikt en bereid gevonden om deel te nemen aan het onderzoek op 29 jaar. De bruto-responsrate op 29 jaar (ten opzichte van de startgroep op 23-jarige leeftijd) bedraagt respectievelijk 55% en 64% voor de geboortecohorte van 1976 en 1980. De belangrijkste verklaring voor het veel groter aantal 29-jarigen bij de laatste bevraging is het ontbreken van de tussenliggende enquête op 26 jaar. Immers, respondenten van de eerste cohorte die niet deelnamen op 26 jaar werden niet meer verzocht deel te nemen aan de enquête op 29 jaar. De bruto-responsrate bij de eerste cohorte uitgesplitst per deelnamegolf geeft immers relatief hoge cijfers, met 68% voor de enquête op 26 jaar ten opzichte van de startgroep van 23-jarigen, en 80% voor de enquête op 29 jaar ten opzichte van de overblijvende groep op 26 jaar.

Wanneer het gaat om representativiteit kunnen ook verschillen worden verwacht tussen de twee (uiteindelijke) steekproeven. Op basis van zowel de duur van het veldwerk als de globale bruto- en netto-responscijfers, én zoals wordt bevestigd in de wegingscoëfficiënten blijkt de meest representatieve steekproef op 23-jarige leeftijd die van cohorte 1980 te zijn, gevolgd door cohorte 1976 (en als laatste cohorte 1978) (Laurijssen, 2005). Wanneer de verdere uitval na 23 jaar in selectiviteit niet verschilt tussen de geboortecohorten, dan nog zal de uiteindelijke steekproef op 29 jaar iets selectiever zijn op bepaalde achtergrondkenmerken voor de geboortecohorte van 1976 dan die van 1980 (zoals zich dat uit in grotere wegingscoëfficiënten). We bekijken hier verder in welke mate de verdere deelname aan het onderzoek, na 23 jaar, meer of minder selectief is geweest bij beide geboortecohorten. Verder vergelijken we de wegingscoëfficiënten op 29-jarige leeftijd, die de vergelijking van een aantal achtergrondkenmerken van de respondenten met die van de populatie (van dezelfde cohorte) reflecteren, en bepaald werden zoals beschreven in Laurijssen (2005).

TABEL 2: GLOBALE VERDELING NAAR GESLACHT, REGIO EN ONDERWIJSNIVEAU OP 23 JAAR, VOOR COHORTE 1976

	Volledige groep op 23		Bereikte groep op 29		Deelname 29 / 23
<b>Geslacht</b>					
Mannen	1544	51,2%	838	50,6%	54,3%
Vrouwen	1471	48,8%	819	49,4%	55,7%
<b>Regio</b>					
Grote stad	332	11,0%	175	10,6%	52,7%
Regionale stad	414	13,7%	210	12,7%	50,7%
Kleine stad	575	19,1%	312	18,8%	54,3%
Niet-stedelijke gemeente	1694	56,2%	960	57,9%	56,7%
<b>Onderwijsniveau (op 23)</b>					
Geen diploma secundair onderwijs	251	8,3%	94	5,7%	37,5%
Beroepssecundair onderwijs	677	22,5%	332	20,0%	49,0%
Technisch secundair onderwijs	517	17,2%	294	17,7%	56,9%
Algemeen secundair onderwijs	474	15,8%	253	15,3%	53,4%
Hoger onderwijs één cyclus	723	24,0%	460	27,8%	63,6%
Universitair onderwijs	365	12,1%	221	13,3%	60,5%

Noot: de gemiddelde deelnamekans 29/23 (= bruto respons-rate) bedraagt 55,0% (uitgesplitst per golf respectievelijk 68,3% en 80,4% op 26 en op 29 jaar)

De kansen op deelname op 29 zijn bij de geboortecohorte van 1976 niet statistisch verschillend voor mannen en vrouwen (ook niet op 26). De deelnamekans neemt globaal wel toe met het onderwijsniveau dat men op 23 jaar bereikt had (zie bovenstaande tabel). De opwaartse trend met een hoger onderwijsdiploma was nog het meest duidelijk op 26. Het zijn vooral de ongekwalificeerden (op 23) die nog eens bijkomende uitval kennen tussen 26 en 29. Naar regio zijn er, in tegenstelling tot op 26 jaar, voor de deelnamekans op 29 wel statistisch significante verschillen. Met name jongeren uit meer stedelijke gebieden (woonplaats op 23) hebben een iets lagere kans om op 29 deel te nemen aan het onderzoek.

Naar achtergrondkenmerken is de deelnamekans bij cohorte 1980 op 29-jarige leeftijd gelijkaardig voor mannen en vrouwen (respectievelijk 64,1 en 64,3% deelnamekans), beduidend kleiner bij respondenten die op 23-jarige leeftijd in een grote stad woonden (deelnamekans van slechts 59,5%), en het grootst bij hogergeschoolden (deelnamekans tot 69,9% bij een diploma van hoger onderwijs van één cyclus, op 23-jarige leeftijd, versus 53,5% voor de laag- en ongekwalificeerden).

TABEL 3: GLOBALE VERDELING NAAR GESLACHT, REGIO EN ONDERWIJSNIVEAU OP 23 JAAR, VOOR COHORTE 1980

	Volledige groep op 23		Bereikte groep op 29		Deelname 29 / 23
<b>Geslacht</b>					
Mannen	1528	51,1%	980	51,0%	64,1%
Vrouwen	1465	48,9%	942	49,0%	64,3%
<b>Regio</b>					
Grote stad	358	12,0%	213	11,1%	59,5%
Regionale stad	479	16,0%	307	16,0%	64,1%
Kleine stad	630	21,0%	414	21,5%	65,7%
Niet-stedelijke gemeente	1526	51,0%	988	51,4%	64,7%
<b>Onderwijsniveau (op 23)</b>					
Geen diploma secundair onderwijs	275	9,2%	147	7,6%	53,5%
Beroepssecundair onderwijs	655	21,9%	412	21,4%	62,9%
Technisch secundair onderwijs	538	18,0%	341	17,7%	63,4%
Algemeen secundair onderwijs	372	12,4%	227	11,8%	61,0%
Hoger onderwijs één cyclus	741	24,8%	518	27,0%	69,9%
Universitair onderwijs	412	13,8%	277	14,4%	67,2%

Noot: de gemiddelde deelnamekans 29/23 (= bruto respons-rate) bedraagt 64,2%.

De vergelijking van beide cohorten van de uiteindelijke uitval in termen van de genoemde achtergrondkenmerken valt niet duidelijk uit in het voordeel van de ene of andere. Beide cohorten hebben een genderneutrale uitval, een grotere uitval bij de jongvolwassenen die in meer stedelijke gebieden woonden, en een grotere uitval bij lagergeschoolden. Het grote verschil in gemiddelde deelnamekans bij beide cohorten bemoeilijkt de vergelijking van de mate van selectieve uitval (we kunnen niet zomaar de responsrates van elkaar aftrekken). Een manier om de relevantie van de selectiviteit op de achtergrondkenmerken te beoordelen, is te kijken naar de grootte van de wegingscoëfficiënten die nodig zijn om de groep van respondenten qua verdeling op de achtergrondkenmerken (meer) te doen aansluiten op die van de populatie (van de geboortecohorte).

TABEL 4: WEGINGSCOËFFICIËNTEN VAN DE TWEE COHORTEN - STATISTIEKEN

	Aantal	Minimum	Maximum	st.afw.
<b>weging c76(23)</b>	3015	0,79	1,82	0,21
<b>weging c76(26)</b>	2060	0,76	2,34	0,30
<b>weging c76(29)</b>	1657	0,71	3,02	0,37
<b>weging c80(23)</b>	2993	0,75	1,66	0,20
<b>weging c80(29)</b>	1922	0,69	1,98	0,27



In de tabel hierboven staan de wegingscoëfficiënten beschreven per cohorte, en per leeftijdsmeting. Daaruit blijkt duidelijk, bij beide geboortecohorten, een toename van de variabiliteit (zowel in de range van de scores als in termen van de standaardafwijking) van de wegingscoëfficiënten met de leeftijd. Dat bevestigt de vastgestelde toename van de selectiviteit op de achtergrondkenmerken met elke bijkomende enquête. De vergelijking van beide cohorten op 29-jarige leeftijd lijkt in het voordeel van de laatste geboortecohorte. Alleen kan die betere match met de populatiegegevens niet louter aan de opvolgenquêteringen worden toegewezen, want reeds op 23-jarige leeftijd is de range (en vooral het maximum, dus de correctie voor ondervertegenwoordiging van bepaalde groepen) veel beperkter. Het is dan ook onduidelijk of in dit opzicht de bevraging meer of minder een relevante rol gespeeld heeft.

## 1.2 Evaluatie van het interview op 29-jarige leeftijd

De effecten van een bevraging meer of minder op de motivatie van de respondenten kan op verschillende manieren uitdraaien. We onderscheiden twee mogelijkheden, en beide leiden tot tegengestelde verwachtingen. Ten eerste kan het zo zijn dat de respondenten van de laatste geboortecohorte op 29 jaar meer gemotiveerd zijn dan die van de eerste cohorte, omdat ze nog niet zo vaak werden geënquêteerd. Ten tweede kan het zijn dat, onder andere door de beperktere - tussentijdse - uitval, ook de minder gemotiveerde respondenten nog in de pool van respondenten zijn overgebleven. Daarnaast kan ook de duur van het interview haar eigen effect hebben in de richting van drukken van de motivatie tot deelname, en in het bijzonder de bereidheid tot latere deelname.

Zoals verwacht, omdat een groter deel van de loopbaan moet geregistreerd worden, duurde de enquête bij c80(29) beduidend langer dan bij c76(29): gemiddeld een uur en 19 minuten, zijnde 13 minuten langer dan bij de eerste cohorte op 29 jaar. Dat de enquête te lang wordt wanneer een groter deel van de loopbaan moet geregistreerd worden, vormt duidelijk een risico, wat ook blijkt uit de reactie bij de respondenten: bij de cohorte van 1980 wordt (veel) vaker geoordeeld dat het interview te lang duurde (34% tegen 17%). Dat reflecteert zich ook in een lagere beoordeling van de mate waarin men het interview aangenaam vond (84% versus 87% van de respondenten vond het interview "aangenaam" of "heel aangenaam"), en in de beperktere toestemming om later opnieuw gecontacteerd te worden (11% weigering tegen 7%). Opvallend is evenwel dat die ontevredenheid met de duurtijd van het interview niet zozeer de beoordeling door de interviewer van de respondent of de beoordeling door de respondent van de zin van dergelijke onderzoeken lijkt te beïnvloeden.

TABEL 5: EVALUATIE DOOR DE RESPONDENT VAN HET INTERVIEW EN EVALUATIE DOOR DE INTERVIEWER VAN DE RESPONDENT, OP 29 JAAR

	1976(29)	1980(29)	sign. vh verschil
Duurtijd interview (uren, minuten)	1u06min	1u19min	***
<i>Evaluatie door respondent</i>			
ervaring interview in het algemeen: aangenaam (1-5)	4,03	3,96	***
interview duurde te lang (in %)	16,9%	34,4%	***
zin van dergelijke onderzoeken (1-5)	3,88	3,85	
onduidelijke of moeilijke vragen (in %)	10,4%	9,3%	
te persoonlijke vragen (in %)	8,3%	9,2%	
mag opnieuw contact worden opgenomen (in %)	93,4%	89,2%	***
<i>Evaluatie door interviewer</i>			
weerstand ondervonden bij rp (in %)	2,8%	3,4%	
beoordeling motivatie respondent (1-5)	4,43	4,44	
beoordeling eerlijk geantwoord (in %)	98,6%	98,1%	
globale beoordeling respondent: gemotiveerd tot antwoorden (1-5)	4,44	4,48	°
N	1657	1922	

Noot met betrekking tot de aantallen: er zijn geen noemenswaardige aantallen missings, wel iets meer bij de vraag naar de zin van dergelijke onderzoeken (respectievelijk 1,1 en 1,5% missings bij cohorte 1976 en 1980).

\*\*\*  $p < 0,001$ ; °  $p < 0,10$

De toestemming om opnieuw gecontacteerd te worden, is zeker bij panelonderzoek van belang. Uit correlaties blijkt echter geen duidelijk verband met de effectieve duurtijd van de enquête. Wel hangt die toestemming samen met de beoordeling van de duur van het interview, van de mate waarin het interview als aangenaam werd ervaren, en met de beoordeling van de zin van het interview door de respondent (deze drie elementen geven in een logistische regressie een Nagelkerke  $R^2$  van 20,3%<sup>1</sup>, zie tabel 6 hieronder). Maar ook gecontroleerd voor deze beoordelingen, blijft de kans op toestemming hoger bij de respondenten van de eerste cohorte dan bij de cohorte 1980.<sup>2</sup>

<sup>1</sup>De verklaarde variantie neemt nog iets toe wanneer de beoordeling van het interview als categorische variabelen wordt opgenomen in de analyse (Nagelkerke  $R^2$  van 20,9%, of bij analyse per cohorte: 20,3% bij 1980 en 21,1% bij 1976).

<sup>2</sup>Die variabelen geven weliswaar een gedeeltelijke verklaring: de relatieve kansverhouding (odds ratio) is na controle nog 'slechts' 1,4 keer groter bij de eerste cohorte in plaats van 1,7 zoals logistische regressies uitwijzen (cijfers verder niet opgenomen).

TABEL 6: LOGISTISCHE REGRESSIE-ANALYSE VAN DE TOESTEMMING OM LATER OPNIEUW GECONTACTEERD TE WORDEN

	<b>B</b>	<b>Exp(B)</b>	<b>Sig.</b>
zin van dergelijke onderzoeken (1-5)	1,03	2,79	***
interview duurde te lang (0/1)	-0,89	0,41	***
ervaring interview in het algemeen: aangenaam (1-5)	0,51	1,66	***
Constante	-3,00	0,05	***

Noot: analyse voor beide datasets samen (N= 3.523). Bij de cohorte van 1980 kunnen de drie variabelen iets meer van de variantie verklaren dan bij de cohorte van 1976, Nagelkerke R<sup>2</sup> respectievelijk 20,7% en 18,6% (maar de resultaten zijn in essentie dezelfde).

\*\*\* p<0,001

De beperktere (vervolg-) contactbereidheid bij de cohorte van 1980 op 29-jarige leeftijd kan te maken hebben met nog andere kenmerken van het interview, naast de duur van het interview, maar kan ook te maken hebben met de selectie van de respondenten. Hiervoor bleek al dat globaal gezien de groep van 29-jarige respondenten van de cohorte 1980 minder selectief blijkt op een aantal achtergrondkenmerken dan de groep 29-jarigen van cohorte 1976. Mogelijk zijn in de pool van c80(29) ook de minder gemotiveerde respondenten nog opgenomen (en niet voordien uitgevallen). Een vergelijking van deze respondenten op het vlak van hun oordeel van het interview op 23-jarige leeftijd kan ons iets meer vertellen over het eventuele selectief behoud van (gemotiveerde) respondenten en intervieweffecten.

### 1.3 Evaluatie van het interview op 23-jarige leeftijd

Al op 23-jarige leeftijd bleek het interview bij de (uiteindelijke respondenten op 29 jaar van) cohorte van 1980 langer dan bij de cohorte van 1976 (zie volgende tabel). Dat kan wellicht toegeschreven worden aan veranderingen in het onderzoeksinstrument, eerder dan dat noodzakelijk de loopbanen zelf in complexiteit toenamen tussen beide cohorten (we denken aan, ondermeer, CAPI in plaats van PAPI, en, daar ook mee verbonden, de meer gestructureerde registratie van de volledige loopbaan, met meer bevroegde arbeidskenmerken, de bijkomende registratie van opleidingen in het kader van de job, ...). Daarmee in overeenstemming vond al een groter aandeel van de respondenten van cohorte 1980 de enquête te lang op 23 jaar. Mogelijk daarmee verbonden werd het interview door de cohorte van 1980 als iets minder aangenaam ervaren, en werd de motivatie van de respondenten van cohorte 1980 als wat minder beoordeeld door de interviewer, dan het geval is bij de geboortecohorte van 1976. Dat de overblijvers van de eerste cohorte evenwel geen meer gemotiveerde selectie van de respondenten vormt, wordt bevestigd door de (tegengestelde) beoordeling door de enquêteur van de eerlijkheid van de antwoorden van de respondenten, die hoger ligt bij de geboortecohorte van 1980.

TABEL 7: EVALUATIE DOOR DE RESPONDENT VAN HET INTERVIEW EN EVALUATIE DOOR DE INTERVIEWER VAN DE RESPONDENT, OP 23 JAAR

	1976(29)	1980(29)	sign. vh verschil
Duurtijd interview (uren, minuten)	55min	1u03min	***
<i>Evaluatie door respondent</i>			
ervaring interview in het algemeen: aangenaam (1-5)	4,02	3,95	***
interview duurde te lang (in %)	7,7%	14,7%	***
zin van dergelijke onderzoeken (1-5)	3,90	3,90	
onduidelijke of moeilijke vragen (in %)	7,5%	8,5%	
te persoonlijke vragen (in %)	6,2%	8,8%	**
<i>Evaluatie door interviewer</i>			
weerstand ondervonden bij rp (in %)	3,1%	2,6%	
beoordeling motivatie respondent (1-5)	4,50	4,38	***
beoordeling eerlijk geantwoord (in %)	95,2%	98,6%	***
globale beoordeling respondent: gemotiveerd tot antwoorden (1-5)	4,46	4,45	
N	1657	1922	

Noot: selectie van de respondenten die op 29-jarige leeftijd werden geïnterviewd.

\*\*\*  $p < 0,001$ ; \*\*  $p < 0,01$

Dat er niettemin een zekere “persoonlijke stabiliteit” is in de beoordeling van de interviews op beide leeftijden, blijkt wel uit de telkens positieve correlatie tussen overeenkomstige kenmerken op beide leeftijden (zie de correlatiecoëfficiënten in de volgende tabel). Respondenten die blijf geven van meer interesse bij het eerste interview, doen dat ook bij het interview op 29-jarige leeftijd, hetzelfde geldt voor hoe zij door de interviewer worden gepercipieerd. Respondenten voor wie het eerste interview langer duurt, hebben ook een langer interview op 29 jaar (of doordat hun loopbanen complexer zijn, of doordat zij er meer de tijd voor nemen).

TABEL 8: CORRELATIECOËFFICIENT VAN DE BEOORDELING VAN HET INTERVIEW EN DE RESPONDENT OP 23 MET DIE OP 29 JAAR

Duurtijd interview (uren, minuten)	0,099	***
<i>Evaluatie door respondent</i>		
ervaring interview in het algemeen: aangenaam (1-5)	0,219	***
interview duurde te lang (in %)	0,160	***
zin van dergelijke onderzoeken (1-5)	0,240	***
onduidelijke of moeilijke vragen (in %)	0,096	***
te persoonlijke vragen (in %)	0,134	***
<i>Evaluatie door interviewer</i>		
weerstand ondervonden bij rp (in %)	0,028	
beoordeling motivatie respondent (1-5)	0,151	***
beoordeling eerlijk geantwoord (in %)	0,002	
globale beoordeling respondent: gemotiveerd tot antwoorden (1-5)	0,124	***

Noot: alle correlaties statistisch significant met  $p < 0,001$ , behalve de beoordeling van de interviewer van eventueel ervaren weerstand en van de eerlijkheid van de antwoorden.

Data: beide datasets samen geanalyseerd,  $N=3.579$  (maar geen grote verschillen tussen beide datasets).

In welke mate de verschillende beoordeling van het interview op 29 jaar door de respondenten van beide cohorten te maken heeft met de selectie van respondenten (in termen van beoordeling van het interview op 23), is op basis van bovenstaande cijfers moeilijk uit te maken. Maar de analyses bevestigen dat dit een rol kan spelen in de (relatief) beperktere bereidheid van de respondenten van c80(29) om nog op latere leeftijd opnieuw gecontacteerd te worden, want ook op 23-jarige leeftijd blijken ze al iets minder gemotiveerd (al dan niet door kenmerken van het interview zelf) dan de respondenten van c76(29).

#### 1.4 Het interview op 23: overblijvers versus uitvallers

Een nog betere toetsing van mogelijke selectieve uitval kan worden gedaan door de "overblijvers" te vergelijken met de "uitvallers" (op 29 jaar) op die evaluatie van de eerste enquête (op 23 jaar).

Het interview op 23-jarige leeftijd geeft reeds goede aanwijzingen over wie op 29-jarige leeftijd zal deelnemen. De deelnemers op 29 jaar vonden reeds op 23 jaar het interview (gemiddeld) aangener, beoordelen de zin van dergelijke onderzoeken iets positiever, en worden door de enquêteurs als meer gemotiveerd, zowel in het algemeen, als om te antwoorden, beoordeeld. Bijkomend vonden de deelnemers op 29 jaar het interview op 23 jaar minder vaak te lang (enkel significant voor cohorte 1980), en oordeelden de interviewers dat de 29-jarigen eerlijker antwoordden op 23 jaar (enkel voor cohorte 1976).

TABEL 9: OVERBLIJVERS VERSUS UITVALLERS, EN DE EVALUATIE VAN HET EERSTE INTERVIEW

	cohort 1976		cohort 1980		
	nt tot 29	tot 29	nt tot 29	tot 29	
Duurtijd interview (uren, minuten)	56min	55min	1u03min	1u03min	
<i>Evaluatie door respondent</i>					
ervaring interview in het algemeen: aangenaam (1-5)	3,97	4,02 *	3,90	3,95 *	
interview duurde te lang (in %)	9,1%	7,7%	19,1%	14,7%	**
zin van dergelijke onderzoeken (1-5)	3,80	3,90 ***	3,76	3,90 ***	
onduidelijke of moeilijke vragen (in %)	8,9%	7,5%	8,5%	8,5%	
te persoonlijke vragen (in %)	7,4%	6,2%	10,3%	8,8%	
<i>Evaluatie door interviewer</i>					
weerstand ondervonden bij rp (in %)	4,5%	3,1% °	3,1%	2,6%	
beoordeling motivatie respondent (1-5)	4,32	4,50 ***	4,24	4,38 ***	
beoordeling eerlijk geantwoord (in %)	92,7%	95,2% **	98,9%	98,6%	
globale beoordeling respondent: gemotiveerd tot antwoorden (1-5)	4,29	4,46 ***	4,30	4,45 ***	
N	1358	1657	1071	1922	

\*\*\* p<0,001; \*\* p<0,01; \* p<0,05; ° p<0,10

Ondanks deze verschillen tussen deelnemers en niet-deelnemers aan het onderzoek op 29 jaar in de evaluatie van het interview op 23 jaar, zien we weinig aanwijzingen voor verschillen tussen beide cohorten op het vlak van selectief behoud van respondenten in dat opzicht.

## 2 Detail van de registratie van de loopbaan

Op basis van de SONAR-data kunnen we goed nagaan welke implicaties een bevraging minder, maar een langere geschiedenis en verder teruggaan in de tijd (geheugeneffecten) mogelijk hebben. Voor die analyse maken we gebruik van de gegevens van de kalenderregistratie van de loopbanen van de respondenten. De kalenderregistratie van de loopbaan omvat de gegevens met betrekking tot deelname aan onderwijs en diverse opleidingen en de arbeidsloopbaan. De verschillende types van registraties waarover we in de zogenaamde kalenders beschikken, staan opgesomd in tabel 10. Bij de registratie van die gegevens is het ook mogelijk om verschillende “activiteiten” of bezigheden te combineren (zoals opleiding volgen en werken).

## 2.1 Aantal periodes

De verwachting is dat de respondenten van de geboortecohorte van 1980 minder registraties hebben met betrekking tot hun loopbaan, omdat ze voor een langere periode moeten terugblikken, en mogelijk ondertussen bepaalde zaken uit het oog verloren of vergeten zijn. We denken niet zozeer dat respondenten helemaal vergeten welke activiteit ze wanneer hadden. Wel dat ze bijvoorbeeld bepaalde details uit het oog verloren zijn (normaal leiden wijzigingen in kenmerken van eenzelfde baan bijvoorbeeld tot een aparte registratie in het kalenderbestand), andere zaken zijn ze misschien wel vergeten omdat ze voor de respondent (achteraf gezien) weinig betekenis hadden (bijvoorbeeld een kortlopende opleiding, of een vroegtijdig afgehaakte studie).

TABEL 10: GEMIDDELD AANTAL REGISTRATIES VAN DE 29-JARIGEN (OVER DE HELE LOOPBAAN BEKEKEN).

	1976	1980	
secundair onderwijs	6,43	6,33	*
hoger onderwijs	2,91	2,85	
deeltijds beroepssecundair onderwijs (DBSO)	0,05	0,09	**
middenstandsopleiding (VIZO)	0,11	0,12	
VDAB-opleiding	0,16	0,17	
opleiding sociale promotie	0,29	0,25	°
andere opleiding	0,34	0,20	***
opleiding in het kader van de job	1,63	2,66	***
periode van niet-werken	2,00	1,97	
betaald werk	4,35	4,07	**
N	1657	1922	

\*\*\* p<0,001; \*\* p<0,01; \* p<0,05; ° p<0,10

De resultaten in bovenstaande tabel stroken niet meteen met de geformuleerde verwachting. Op twee zaken blijken de 29-jarige respondenten van de cohorte van 1980 het zelfs beter te doen: ze hebben meer registraties van het DBSO, en een groter aantal geregistreerde opleidingen in het kader van het betaald werk. Nu is dat laatste niet verwonderlijk, gegeven dat die opleidingen bij die cohorte van bij het begin werden bevraagd, terwijl bij de cohorte 1976 dit pas werd toegevoegd aan de vragenlijst bij de enquête op 29 jaar. Mogelijk heeft de toename van het aantal geregistreerde DBSO-opleidingen eveneens te maken met een verbetering van het instrument (in het bijzonder omdat de vraag naar DBSO-opleidingen, net zoals voor de VIZO-opleidingen trouwens, duidelijker gesteld wordt in de CAPI-versie).

Andere verschillen in bovenstaande tabel wijzen er wel op dat de laatste cohorte minder registraties heeft voor wat betreft het secundair onderwijs, de opleidingen sociale promotie, andere opleidingen, en minder werkregistraties. Hier gaat het mogelijk om problemen van

herinnering, omdat een grotere periode van de loopbaan op dat vlak bevraagd werd op 29-jarige leeftijd. Of het inderdaad daarmee te maken heeft, kunnen we meer nagaan en beter interpreteren wanneer we het aantal registraties in de kalender uitsplitsen naar die die betrekking hebben op de periode sinds de enquête op 23 jaar, en die die betrekking hebben op de eerste periode, zoals we doen in de volgende tabel.

TABEL 11: GEMIDDELD AANTAL REGISTRATIES VAN DE 29-JARIGEN, OVER DE LOOPBAAN GEMETEN APART VOOR DE LEEFTIJD VAN 23 EN NA 23

	Registraties tot op 23			Registraties 23-29		
	1976	1980		1976	1980	
secundair onderwijs	6,41	6,33	*	0,01	0,00	°
hoger onderwijs	2,42	2,40		0,49	0,45	
deeltijds beroepssecundair onderwijs (DBSO)	0,04	0,09	**	0,01	0,00	°
middenstandsopleiding (VIZO)	0,08	0,10		0,03	0,02	*
VDAB-opleiding	0,06	0,08	*	0,10	0,08	
opleiding sociale promotie	0,07	0,09	*	0,22	0,16	***
andere opleiding	0,19	0,10	***	0,16	0,10	***
opleiding in het kader van de job	0,00	0,64	***	1,63	2,02	***
periode van niet-werken	1,16	1,20		0,84	0,78	
betaald werk	1,51	1,28	***	2,84	2,79	
N	1657	1922		1657	1922	

\*\*\* p<0,001; \*\* p<0,01; \* p<0,05; ° p<0,10

Uit de opsplitsing van de telling van het aantal registraties naar de periode tot de eerste enquête en de periode daarna tot op 29-jarige leeftijd, kunnen we drie zaken afleiden. Ten eerste blijkt de toename van het aantal registraties van DBSO en van de opleidingen in het kader van de job bij de cohorte van 1980 (in vergelijking met die van 1976) reeds duidelijk op 23-jarige leeftijd. Daarmee bevestigen de cijfers de suggestie hiervoor dat de toename in die registraties te maken heeft met verbeteringen in het gebruikte instrument (met betrekking tot opleiding in het kader van de job is het zo dat deze bij cohorte 1976 zelfs pas werden bevraagd bij de enquête op 29 jaar). Een tweede vaststelling is dat voor wat betreft de andere verschillen tussen beide cohorten, de lagere aantallen registraties bij de tweede cohorte inderdaad bevestigd worden voor het deel van de registratie van de loopbaan tussen de leeftijd van 23 en 29. Respondenten van de laatste cohorte registreerden tussen beide leeftijden een lager aantal periodes van secundair onderwijs, opleiding sociale promotie, andere opleiding (en betaald werk, maar ns), dan de respondenten van de eerste cohorte. Een derde vaststelling is dat voor de meeste van die verschillen in aantallen tussen beide cohortes evenwel ook geldt dat reeds tot op 23-leeftijd blijkt dat er sprake is van een lager



aantal registraties bij de cohorte van 1980. De enige uitzondering daarop blijkt het aantal opleidingen sociale promotie te zijn.

Vanwaar die grote verschillen (reeds) op 23-jarige leeftijd? We overlopen twee mogelijkheden: ofwel spelen veranderingen in het instrument een rol, ofwel heeft het met andere kenmerken van de respondenten (op 29 jaar) te maken.<sup>3</sup> Met betrekking tot het instrument, zijn heel wat veranderingen mogelijk doordat de bevraging op 23-jarige leeftijd bij de eerste cohorte nog op papier verliep (PAPI), terwijl de respondenten bij alle andere enquêtes een geprogrammeerde vragenlijst doorliepen (CAPI). Maar bij deze verklaring zouden we minder verschillen verwachten voor de periode die volgt op de enquête op 23 jaar. Dat het aan het instrument ligt lijkt daarom geen afdoende verklaring voor de cohorteverschillen. Anderzijds kan het zijn dat de (overblijvende) groep van respondenten over beide cohorten verschilt doordat bij de laatste cohorte een grotere groep tot op de leeftijd van 29 jaar deelneemt (door het ontbreken van een bevraging op 26 jaar, waarop bijkomende uitval mogelijk is). Daarbij is dan de vraag of die grotere groep die behouden blijft, ook betekent dat het om een minder selectieve groep gaat (in termen van relevante loopbaankenmerken). Om daarover meer te kunnen zeggen, vergelijken we hierna het aantal registraties van de beide cohorten niet voor de selectie van 29-jarigen, maar wel op basis van alle oorspronkelijke respondenten (de “23-jarigen”).

TABEL 12: VERGELIJKING VAN HET GEMIDDELD AANTAL REGISTRATIES VOOR DE LEEFTIJD VAN 23 (SELECTIE: ALLE RESPONDENTEN)

	cohorte 1976	cohorte 1980	
secundair onderwijs	6,39	6,32	*
hoger onderwijs	2,27	2,34	
deeltijds beroepssecundair onderwijs (DBSO)	0,07	0,10	**
middenstandsopleiding (VIZO)	0,09	0,11	
VDAB-opleiding	0,06	0,08	*
opleiding sociale promotie	0,07	0,09	**
andere opleiding	0,18	0,12	***
opleiding in het kader van de job	0,00	0,60	***
periode van niet-werken	1,25	1,30	
betaald werk	1,83	1,54	***
N	3007	2993	

\*\*\* p<0,001; \*\* p<0,01; \* p<0,05

<sup>3</sup>Een alternatieve verklaring kan zijn dat de loopbanen van jongvolwassenen in de periode tussen beide geboortecohorten gewijzigd zijn. Dan is het onmogelijk om stevige conclusies te trekken door de vergelijking van beide cohorten over het eventuele effect van een bevraging meer of minder.

In bovenstaande tabel komen dezelfde verschillen naar voor als in (de eerste kolommen van) de vorige tabel (tabel 11). Het gemiddeld aantal registraties (tot op de leeftijd van 23) ligt hoger bij de respondenten van de geboortecohorte 1980 voor wat betreft het deeltijds beroepssecundair onderwijs, de VDAB-opleidingen en de opleidingen voor sociale promotie. Het gemiddeld aantal registraties ligt lager voor wat betreft secundair onderwijs, andere opleiding, en betaald werk. Het beperkter aantal ‘andere opleidingen’ heeft mogelijk te maken met een wijziging van het onderzoeksinstrument, nl. de aparte registratie van de opleidingen in het kader van de job bij de laatste cohorte (bij de eerste cohorte was hiervoor geen aparte registratie voorzien op 23-jarige leeftijd, en dan werden bepaalde opleidingen in het kader van de job wellicht geregistreerd bij de ‘andere opleidingen’). De verschillen tussen beide cohorten zijn doorgaans ook van dezelfde grootte-orde bij vergelijking van alle 23-jarigen met de 29-jarigen (ook al liggen de absolute cijfers voor betaald werk bijvoorbeeld hoger), of iets beperkter (eventueel voor de registraties van secundair onderwijs en andere opleiding). Zo lijkt te worden bevestigd dat inderdaad selectie op een aantal kenmerken plaatsvindt.

TABEL 13: VERGELIJKING VAN HET GEMIDDELD AANTAL REGISTRATIES VOOR DE LEEFTIJD VAN 23 VAN DE RESPONDENTEN DIE TOT 29 WERDEN BEVRAAGD (“OVERBLIJVERS”), EN DE ANDERE RESPONDENTEN (“UITVALLERS”)

	cohorte 1976			cohorte 1980		
	uit- vallers	over- blijvers		uit- vallers	over- blijvers	
secundair onderwijs	6,35	6,43	°	6,31	6,33	
hoger onderwijs	2,07	2,43	***	2,22	2,40	*
deeltijds beroepssecundair onderwijs (DBSO)	0,09	0,04	***	0,12	0,09	*
middenstandsopleiding (VIZO)	0,10	0,08		0,13	0,10	
VDAB-opleiding	0,07	0,06		0,08	0,08	
opleiding sociale promotie	0,07	0,07		0,10	0,09	
andere opleiding	0,18	0,19		0,15	0,10	**
opleiding in het kader van de job	0,00	0,00		0,54	0,64	°
periode van niet-werken	1,37	1,16	***	1,47	1,20	***
betaald werk	2,21	1,52	***	2,01	1,28	***
N	1353	1654		1071	1922	

\*\*\* p<0,001; \*\* p<0,01; \* p<0,05; ° p<0,10

Een meer doorgedreven vergelijking van de eventuele selectieve respons van de 29-jarigen kan op basis van de cijfers in bovenstaande tabel. We vergelijken voor beide cohorten het gemiddeld aantal registraties tussen de respondenten die tot op 29-jarige leeftijd gevolgd en bevroegd werden (de “overblijvers”) en de groep van respondenten die op dat tijdstip niet meer bevroegd werd (de “uitvallers”). Dan blijken er inderdaad behoorlijk wat verschillen die wijzen op een selectieve paneluitval. Zo is het aantal registraties hoger onderwijs groter bij

de 29-jarigen, het aantal registraties DBSO kleiner, zijn er voor de “overblijvers” minder periodes van niet-werken én van betaald werk. Die verschillen zijn consistent met een patroon waarbij hogergeschoolde respondenten (gemakkelijker) in de steekproef blijven tot op het latere meetmoment (zoals ook bleek uit de analyse van de representativiteit van de steekproef in het eerste deel van de paper): ze studeren langer, ze hebben daardoor op 23 jaar een kortere arbeidsloopbaan, maar ook, door hun hoger onderwijsniveau, (gemiddeld) een minder problematische arbeidsmarktintrede (en dus minder periodes van werkloosheid of andere inactiviteit).

Voor onze vergelijking hier, is de relevante vraag evenwel of de selectiviteit van de paneluitval anders is tussen de beide geboortecohorten (naar aanleiding van een meetmoment minder, en gegeven de vaststelling van een grotere overblijvende groep bij de cohorte van 1980). Op basis van bovenstaande cijfers zijn die verschillen, als ze er al zijn, in elk geval niet groot. Op het vlak van het verschil in het aantal registraties van het hoger onderwijs is de kloof tussen de overblijvers en uitvallers misschien wat kleiner bij de laatste cohorte. Op het vlak van de registratie van het aantal opleidingen, compenseert het verschil in ‘andere opleidingen’ en in ‘opleidingen in het kader van de job’ elkaar wellicht. Met betrekking tot het aantal registraties van niet-werken en van betaald werk is de kloof tussen de overblijvers en de uitvallers daarentegen even groot in beide cohorten. Mogelijk is er, zoals ook de analyse van de representativiteit van de uiteindelijke steekproef leek uit te wijzen, sprake van een afname van de selectiviteit in de overblijvende respondenten, al zijn de aanwijzingen daarvoor eerder beperkt.

Die beperktere selectiviteit biedt mogelijk ook een verklaring voor het beperkter aantal registraties secundair onderwijs bij c80(29) (maar dan is het wel verrassend dat we geen verschillen vinden tussen de 29-jarigen van beide cohorten in het aantal registraties hoger onderwijs). We weten dus niet zeker in welke mate het lager aantal registraties secundair onderwijs het gevolg is van een bevraging minder. Zoals reeds aangehaald vormt de categorie ‘andere opleidingen’ een bijzondere, een die eventueel in combinatie met de opleidingen in het kader van de job moet worden beschouwd, en waarvoor dan weinig conclusies mogelijk zijn. Met betrekking tot het (beperkter) aantal werkregistraties is er (ook) geen duidelijk effect van een bevraging meer of minder, omdat enkel voor de leeftijd van 23 jaar de laatste cohorte minder registraties in verband met betaald werk heeft (wat suggereert dat in hoofdzaak wijzigingen in het instrument aan de basis liggen van het kleiner aantal werkregistraties bij de laatste cohorte). Het meest duidelijk zijn de verschillen tussen beide cohorten op het vlak van opleidingen voor sociale promotie, waar de registratie door de 29-jarige respondenten van de tweede cohorte minder goed (b)lijkt (en niet voor de leeftijd van 23). De mindere kwaliteit van de bevraging bij de tweede cohorte op dat vlak kan niet worden toegewezen aan veranderingen in selectiviteit (er blijken immers geen

verschillen in overblijvers versus uitvallers). Maar al bij al gaat het toch om relatief beperkte (kwantitatieve) verschillen.

## 2.2 Missing data

Bij de SONAR-enquêtes is er weinig missing data, omdat de geprogrammeerde vragenlijst niet toelaat om bijvoorbeeld vragen over te slaan. Het grootste probleem van missing data bij de CAPI-bevragingen is dat bij een foutief antwoord op een vraag met belangrijke filtergevolgen bepaalde vragen (onterecht) niet of wel gesteld worden. En die fouten kunnen soms achteraf, bijvoorbeeld aan de hand van open antwoorden, wel enigszins rechtgezet worden, maar de ontbrekende informatie kan zelden helemaal gerecupereerd worden. De manier waarop getracht wordt dit soort fouten zo veel mogelijk te beperken, is door een doorgedreven opleiding van de enquêteurs, en door verfijning van het instrument (onder meer door extra toelichting bij de belangrijke filtervragen). Als een enquêteur zelf merkt dat het foute antwoord werd ingegeven (omdat het vervolg niet het verwachte vervolg is), is het het wel mogelijk om terug te keren in de vragenlijst.

Een mogelijke toetsing van de hoeveelheid informatie die een respondent nog geeft wanneer een langere loopbaan moet worden gereconstrueerd, kunnen we bijvoorbeeld wel nagaan op open vragen die nadien gecodeerd werden. We bekijken hier de aangeleverde informatie met betrekking tot de codering van de beroepen. Het aantal beroepen dat bevraagd wordt tijdens het interview, is afhankelijk van het aantal keren dat een respondent van functie verandert, of werkonderbrekingen heeft in de arbeidsloopbaan. Voor elk van die beroepen dient de respondent (uitgebreid) te beschrijven welke taken en handelingen het werk inhoudt. Vervolgens werden de beschrijvende gegevens telkens gecodeerd. Hoewel naast eventuele geheugen- of inspanningseffecten verschillen tussen cohorten ook beïnvloed zouden kunnen zijn door de vaardigheid van de codeurs, door eventuele andere gemiste filters, enz., maakten we een vergelijking van het aantal beroepen dat niet of onvoldoende kon worden gecodeerd. Uit die vergelijking blijkt geenszins dat de respondenten van c80(29) minder goed hun beroep omschreven. Zo zijn er in het coderingsbestand slechts 2 gevallen gedocumenteerd waar er onvoldoende informatie was om een SBC-code te kunnen toekennen voor de enquête bij cohorte 1980 op 29 jaar, tegen respectievelijk 18 en 27 op 26- en 29-jarige leeftijd bij de cohorte van 1976.

## 2.3 Afronding van gegevens

Als laatste element in de evaluatie van de effecten van een (tussentijdse) bevraging meer of minder, bekijken we de mate waarin de periodes van de kalenderregistraties “afgerond”

werden. Het gaat dan concreet over de start- en einddata die respondenten opgeven voor elk van de periodes. Wanneer bepaalde gebeurtenissen (zoals het starten in een baan) op volledig toevallige wijze verdeeld zouden zijn over het jaar, dan kan verwacht worden dat elke maand even vaak voorkomt (als element in de startdatum, of in de einddatum). Wanneer bepaalde maanden er dan uit springen (in relatief voorkomen, meer dan 1/12) dan wijst dat op een systematiek in de periodes. En uiteraard zijn de onderwijsperiodes, periodes van opleidingen, en periodes van betaald werk niet random verdeeld over een jaar. Zo volgen onderwijsperiodes die van schooljaren, zijn periodes van opleidingen eventueel gestructureerd in (kalender)jaren, en kunnen periodes van betaald werk bijvoorbeeld iets vaker aanvatten in januari, of rond de zomer, dit laatste in het bijzonder voor schoolverlaters. Maar wanneer geregistreerde periodes meer systematiek laten zien over verschillende cohorten, dan kan vermoed worden dat sprake is van "afronding" van de data. We verwachten dat wanneer geheugeneffecten een grotere rol (kunnen) spelen (i.c. cohorte 1980, in vergelijking met cohorte 1976), dat bepaalde maanden vaker zullen voorkomen, in het bijzonder deze die wijzen op een zekere "afronding" of "normalisering". Zo verwachten we voor schoolperiodes dat er minder afwijkende start/einddata zullen zijn, en vaker typisch volledige schooljaren geregistreerd zullen worden, wanneer respondenten verder in de tijd terug moeten blikken voor de rapportering van de periode.<sup>4</sup> Met betrekking tot de andere types van registraties (opleidingen, betaald werk) verwachten we dat januari vaker als startmaand, en december vaker als eindmaand voorkomt indien geheugeneffecten een rol spelen.

We toetsen deze verwachtingen voor het verschil tussen beide cohorten. Schooljaren bakenen we af van september tot juni (waarover enquêteurs ook heel sterk gebriefd worden), en we tellen het aantal periodes met dergelijke "standaard" start- en einddatum. Voor de andere periodes tellen we het aantal periodes met een "standaard" start op januari en/of einde op december (zie onderstaande tabel voor de cijfers).

Het meest opvallend verschil in verandering in de "standaardisering" van de periodes is er voor de registratie van de VIZO-opleidingen. De sterke toename over de cohorten van de rapportering van VIZO-opleidingen in termen van schooljaren heeft waarschijnlijk te maken met een instrumentwijziging (zie ook verder). Daarnaast blijkt dat de tweede cohorte minder standaard registraties heeft voor de VDAB-opleidingen ( $p < 0,05$ ), en ook voor betaald werk (dit laatste slechts significant op 0,10 niveau, terwijl het toch om een groot aantal registraties gaat, nl. 15.026). Aparte toetsen voor de startmaand en eindmaand geven aan dat voor de VDAB-opleidingen het verschil niet zozeer zit in de beginmaand, wel in de eindmaand, die in de tweede cohorte minder vaak december is (6% tegen 11% in de eerste cohorte,  $p < 0,05$ ).

---

<sup>4</sup>De flexibilisering van het hoger onderwijs, waarvan de effecten tegen deze verwachting in zouden werken, is nog niet (ten volle) tot uiting gekomen bij de hier onderzochte geboortecohorten.

Voor de periodes van werk is er geen statistisch significant verschil voor de beginmaand of de eindmaand apart getoetst.

TABEL 14: AANDEEL "STANDAARD" REGISTRATIES VAN DE 29-JARIGEN PER COHORTE (OVER DE HELE LOOPBAAN GEMETEN)

	1976	1980	sign.verschil
secundair onderwijs	1,00	1,00	
hoger onderwijs	0,97	0,97	
deeltijds beroepssecundair onderwijs (DBSO)	0,95	0,96	
middenstandsopleiding (VIZO)	0,73	0,92	***
VDAB-opleiding	0,21	0,14	*
opleiding sociale promotie	0,15	0,13	
andere opleiding	0,16	0,13	
betaald werk	0,20	0,19	°

\*\*\*  $p < 0,001$ ; \*  $p < 0,05$ ; °  $p < 0,10$

Noot: N verschilt bij elk van de testen (gaande van 247 registraties voor DBSO tot 22.819 registraties voor secundair onderwijs).

Noot: standaard periodes gedefinieerd als: van september tot juni voor de vier types van onderwijs, en voor de opleidingen en het betaald werk: periodes startend op januari en/of eindigend in december.

Deze resultaten geven aan dat er relatief beperkte verschillen zijn tussen de cohorten, en als er verschillen zijn (VDAB, betaald werk) dan wijzen ze eerder op minder standaardisering van de antwoorden (in termen van afronding van de periodes) bij de laatste cohorte dan de eerste cohorte (tenzij met betrekking tot de VIZO-opleidingen, zie ook verdere analyse en berekening).

In volgende tabel splitsen we uit naar de registraties die betrekking hebben op de situatie nog voor de eerste enquête op 23 jaar, en al de daaropvolgende registraties tot 29 jaar.

Opnieuw apart getoetst voor de begin- en einddatum van de geregistreerde periodes, geldt dat het verschil bij de VDAB-opleidingen enkel tot uiting komt in de einddatum ( $p < 0,05$ ). Tussen 23 en 29 jaar werden door de respondenten van cohorte 1980 beduidend minder gestandaardiseerde (einddata van) periodes gerapporteerd voor de VDAB-opleidingen.

Wat betreft de VIZO-opleidingen blijkt de sterkere standaardisering van de periodes bij de tweede cohorte enkel in de registraties tot 23 jaar. Dat is enerzijds uiteraard de periode waarin de meeste van dergelijke opleidingen worden gevolgd. Maar anderzijds is die toename van rapportering in termen van schooljaren bij de vergelijking op 23, en niet (meer) na 23, ook volledig in overeenstemming met een sterkere klemtoon daarop in de vragenlijst en bij de briefing van de enquêteurs, vanaf de introductie van de CAPI-bevraging.

TABEL 15: AANDEEL "STANDAARD" REGISTRATIES VAN DE 29-JARIGEN PER COHORTE, GEMETEN OVER DE LOOPBAAN TOT 23, EN VAN 23 TOT 29 JAAR

	tot 23 jaar			van 23 tot 29 jaar		
	1976	1980	sign.	1976	1980	sign.
secundair onderwijs	1,00	1,00		0,78	1,00	
hoger onderwijs	0,98	0,97	*	0,91	0,92	
deeltijds beroepssecundair onderwijs (DBSO)	0,95	0,96		1,00	-	
middenstandsopleiding (VIZO)	0,73	0,95	***	0,72	0,69	
VDAB-opleiding	0,18	0,17		0,22	0,11	**
opleiding sociale promotie	0,16	0,15		0,15	0,11	
andere opleiding	0,20	0,15		0,13	0,10	
betaald werk	0,20	0,17	*	0,20	0,20	

\*\*\* p<0,001; \*\* p<0,01; \* p<0,05

Nergens zijn hier, na opsplitsing van de periodes, nog aanwijzingen voor meer standaardisering van de rapportering van de periodes wanneer een groter deel van de loopbaan moet worden overschouwd (registratie tussen 23 en 29 jaar). De meeste van de verschillen zijn er al op 23 jaar, en wellicht het gevolg van instrumentwijziging (PAPI/CAPI); en bovendien wijzen de verschillen eerder op minder standaardisering (tenzij dan de VIZO-opleidingen; maar daar is het verschil er al op 23 en heeft het wellicht te maken met de duidelijkere structurering van de registraties en briefing van de enquêteurs in verband met de schooljaren).

### 3 Conclusie

De onderzoeksgroep SONAR heeft zich veel moeite getroost om de eventuele gevolgen van "een bevraging minder" te kunnen opvangen en de respons te maximaliseren voor c80(29). Cijfers in dit rapport illustreren dat de inspanningen hebben geredeneerd: we vinden weinig tot geen aanwijzingen in de data dat de loopbaanregistraties van mindere kwaliteit zouden zijn. De uiteindelijk overgebleven groep respondenten op 29 jaar blijkt minder selectief op een aantal achtergrondkenmerken (dit kan evenwel niet worden toegewezen als gevolg van een bevraging minder). Van alle onderzochte verschillen, kan enkel voor de registratie van opleidingen van sociale promotie geconcludeerd worden dat er verlies aan informatie lijkt te zijn door de langere periode waarvoor respondenten dienden te antwoorden. Voor veel zaken bleken de verschillen beperkt (aantal geregistreerde periodes, missing data, afronding van data), toe te wijzen aan de beperktere selectiviteit, of zelfs verbeterd door ingevoerde instrumentverbeteringen.

Het enige minpunt dat in dit rapport duidelijk naar voren komt, is dat “een bevraging minder” het risico inhoudt dat het afnemen en beantwoorden van de vragenlijst te lang gaat duren (zowel objectief als subjectief), en het eventuele probleem van beperktere bereidheid tot (latere) vervolgdeelname aan het onderzoek. Maar daarbij dient benadrukt, dat er ook aanwijzingen zijn dat die beperktere toestemming voor latere contactname daarom niet eenvoudigweg als het gevolg van de langere duurtijd kan worden geïnterpreteerd.

Tot slot lijken de conclusies in dit rapport aan te geven dat het uitvoeren van een bijkomende (tussenliggende) bevraging niet nodig is, omdat ze weinig bijdraagt tot de betrouwbare reconstructie van de loopbanen, en mogelijk zelfs het maximaal behoud van de steekproef en motivatie van de deelnemers beperkt. Toch kan dit op basis van deze technische analyses niet besloten worden. Een bevraging minder impliceert immers ook dat er een meetmoment minder is waarop de respondent kan worden bevraagd over meer subjectieve inschattingen van de eigen (beroeps)positie, houdingen en waarden. Dat soort van gegevens is zowat onmogelijk retrospectief te bevragen, en evenmin via eventuele alternatieve dataverzameling op basis van administratieve data beschikbaar. In dat opzicht, ook al wordt de feitelijke loopbaan goed geregistreerd, beschikken we over veel minder analysepotentieel op basis van de data met de twee enquêtes van de laatste geboortecohorte dan met de drie bevragingen van de eerste geboortecohorte.