

RAPPORT

VALGUNDERSØGELSEN 2007

*Metoderapport for undersøgelse på TNS
Gallups internetpanel*

Projektperiode:
Januar-maj 2008

Rapporteringsmåned:
Maj 2008

Kunde:
Aalborg Universitet, Institut for
Økonomi, Politik og Forvaltning
Fibigerstræde 1
9220 Aalborg Ø

INDHOLD

1	INDLEDNING	2
2	INTERVIEWTIDSPUNKT	2
3	POPULATION	2
4	STIKPRØVEUDVÆLGELSE.....	2
5	GENNEMFØRELSESSTATISTIK	3
6	BORTFALD	5
7	METODEFORSKELLE OG HOLDNINGSSPØRGSMÅL	8
8	LITTERATUR	10

1 Indledning

Valgundersøgelsen 2007 gennemføres som en kombineret CAPI/CAWI/ Postal undersøgelse. I tillæg er der blevet gennemført en undersøgelse på Gallups internetpanel Gallup@Forum.

Denne metoderapport redegør for anvendelsen af Gallups internetpanel Gallup@Forum herunder undersøgelsens metode og gennemførelsesgrad for skabe grundlag for at vurdere datakvalitet og anvendeligheden.

2 Interviewtidspunkt

Der blev i alt gennemført 2.576 interviews. Langt hovedparten blev gennemført i perioden 25. januar – 6. februar 2008, mens de sidste ca. 250 interviews blev gennemført fra d. 19.-25. februar. Interviewtidspunktet fremgår af datafilen, men man skal dog være opmærksom på, at der kan forekomme udsving i svartiderne. Årsagen er, at når en respondent har påbegyndt besvarelsen begynder tiden at tælle og stopper ikke selvom respondenter lukker ned og genoptager på et senere tidspunkt.

3 Population

Undersøgelsens population er defineret som herboende personer på mindst 18 år med bopæl i en privat husstand og med stemmeret.

4 Stikprøveudvælgelse

Undersøgelsen er gennemført på Gallups internetpanel Gallup@Forum.

Gallup@Forum er et internetpanel, der omfatter ca. 45.000 danskere i alderen 15+. Deltagere i panelet rekrutteres løbende ved de landsrepræsentative CAPI og CATI undersøgelser, som Gallup gennemfører. Det vil sige, at ingen har mulighed for på eget initiativ at tilmelde sig til panelet, da rekrutteringen udelukkende foregår gennem landsrepræsentative CAPI og CATI undersøgelser.

Målsætningen med survey undersøgelser er i hovedreglen at generalisere undersøgelsens resultater til den ønskede population. For at det er muligt, og for at gængse statistiske usikkerhedsberegninger er gyldige, er der en forudsætning om, at stikprøven skal være simpelt tilfældig udvalgt fra populationen. Denne antagelse er dog sjældent mulig at imødekomme helt i praksis. Derfor må man ofte nøjes med en tilnærmelsesvist tilfældig stikprøve (Fx klyngeudvælgelse ved CAPI undersøgelser, cpr-udtræk samt random digit dialing ved CATI undersøgelser). Gallup@Forum udgør kun et lille udsnit af hele den danske befolkning, hvilket betyder, at der ikke er tale om simpel tilfældig udvælgelse fra hele den population (danskere 18+ med stemmeret). Selvom panelet er nøje sammensat er der i selve basen en underrepræsentation af helt unge under 21 (især mænd) og ældre (70+). Endvidere er andelen af befolkningen med en højere uddannelse overrepræsenteret i panelet.

For at opveje for de mindre demografiske skævheder i panelet er stikprøven udtrukket stratificeret efter køn * alder, region og uddannelse blandt de 45.000 i panelet. Det vil sige den

brutto sample, der er udvalgt og som har modtaget en indbydelse til at deltage i undersøgelsen matcher de danske vælgeres sammensætning og ikke selve panelets sammensætning.

Udover demografiske skævheder kan der være andre årsager til at man ikke umiddelbart kan sammenligne CAPI og undersøgelser gennemført på et internetpanel. En nærliggende tanke er, at man som medlem af et internetpanel kan formodes at være mere interesseret i at deltage i spørgeskemaundersøgelser end den gennemsnitlige dansker (og måske også den gennemsnitlige deltager i spørgeskemaundersøgelser). Dette vil også kunne influere respondenternes adfærd i andre sammenhænge, som f.eks. politisk adfærd. Derfor er det relevant at sammenligne respondenterne i denne undersøgelses politiske interesse med resultaterne fra tidligere Valgundersøgelser. I Tabel 2.3 fremgår det da også, at deltagerne i denne undersøgelse er mere politiske interesserede end deltagerne i tidligere undersøgelser. Der er før undersøgelsen udtrukket en stratificeret stikprøve for at korrigere for panelets skævheder. Efter undersøgelsens afslutning er data blevet vejret (køn*alder, region, uddannelse, partivalg 2007) for at korrigere for eventuelle skævheder i besvarelsesprocenter mellem forskellige grupper. En væsentlig forskel mellem stratificering og vejning er dog at partivalg er medtaget som en parameter i vejningen. Det er ikke muligt at stratificere på partivalg i TNS Gallups internetpanel på nuværende tidspunkt.

Forskellene mellem undersøgelserne eksisterer stadig når data er vejret. Med mindre der er sket et skred i danskernes politiske interesse de senere år er panelisterne mere interesserede end de som deltager i CAPI og CATI undersøgelser, hvilket ikke er overraskende. Man kunne forestille sig at forskellene i politisk interesse kan spille ind på andre forhold, som andelen af sofavælgere og lign.

5 Gennemførelsesstatistik

Gennemførelsesstatistikken er en opgørelse over hvor mange interview der er gennemført

Respondenter der af den ene eller anden årsag ikke kan interviewes kaldes under ét for frafald. Der kan sondres mellem neutralt frafald og det egentlige frafald. Det førstnævnte er fx respondenter, som falder uden for målgruppen, eller som ikke kan deltage på grund af tekniske fejl. Det egentlige frafald er det problematiske, det er først og fremmest nægterne, respondenter der afbryder interviewet undervejs, og respondenter der slet ikke har reageret på henvendelsen.

Nedenfor ses gennemførelsesstatistikken for valgundersøgelsen på Gallups internetpanel. Som det fremgår, er stikprøveundersøgelsen baseret på et udtræk på i alt 6.100 respondenter. Heraf er der et mindre, neutralt frafald grundet tekniske fejl og at respondenterne faldt uden for målgruppen (fx ikke stemmeberettiget).

Når det neutrale bortfald fraregnes fremkommer en bruttostikprøve på 6.007 respondenter, svarende til 98,5 pct. af det oprindelige udtræk. Af disse er det lykkedes at gennemføre interview med 2.576 personer eller 42,9 pct. – hvilket er altså undersøgelsens gennemførelsesprocent. Dette er en relativ lav svarprocent for undersøgelser gennemført på Gallups internetpanel.

Antal udsendte weblinks til skema	6100	(i%)
Uden for målgruppen (Fx under 18, ikke stemmeret)	44	
Tekniske fejl (normalt korte udfald på serveren eller lignende) ¹	49	
Netto sample.....	<u>6.007</u>	100,00 %
Nægtet deltagelse	804	13,38 %
Ingen respons på udsending	1.850	30,80 %
Antal afbrudte interview	<u>777</u>	12,93 %
Gennemførte interviews	<u>2.576</u>	42,88%

I forbindelse med gennemførelse af undersøgelsen blev der udsendt to påmindelser til respondenter, som endnu ikke havde svaret. Påmindelserne blev udsendt per E-mail til de respondenter, som enten ikke havde begyndt besvarelsen eller endnu ikke var færdige med at besvare skemaet. Påmindelserne blev udsendt d.1. februar og d. 5. februar.

En lav gennemførelsesprocent betyder dog ikke *nødvendigvis* at undersøgelsen er fejlbehæftet, men at der er *mulighed* for fejl grundet systematisk frafald. Det kræver en nærmere analyse af repræsentativiteten at belyse omfanget af eventuelle fejl. I et vist omfang kan der korrigeres for systematiske fejl ved en efterfølgende vejning. Der er derfor inkluderet en vægt i datafilen. Vægten har til formål at korrigere for nogle af de forskelle som findes mellem populationens karakteristika og deltagerne i undersøgelsen. I afsnittet nedenfor fremgår nogle af de skævheder, som vægten korrigerer for. Med mindre andet fremgår, er der tale om de uvejede data.

¹ Ved så store undersøgelser som denne sender TNS Gallup altid ud til et mindre udsnit af stikprøven, for at sikre, at der ikke er fejl før der sendes ud til alle. Ved gennemførelsen af denne undersøgelse blev der sendt ud til 100 respondenter, hvoraf 22 blev sorteret fra på grund af en teknisk fejl i skemaet. Fejlen blev øjeblikkeligt rettet og der blev sendt ud til de resterende 6.000.

6 Bortfald

Opgørelserne for Valgundersøgelsen G@F i tabel 2.1, 2.2, 2.3 og 2.4 er alle uvægtede.

Tabel 2.1 Svarfordelingen efter køn. Stikprøvepopulation i Valgundersøgelserne G@F i forhold til den samlede befolkning 01.01.08.

	Valgundersøgelse G@F (i %)	Hele befolkningen* (i %)	Hele G@F
Andel af mænd	51,9	49,1	46,0
Andel af kvinder	48,1	51,0	54,0
I alt	100	100	100

*Andel af den danske befolkning 18+, Danmarks Statistik

I Valgundersøgelsen 2008 er N=2576; I hele G@F er N=49.846

Tabel 2.2 Svarfordelingen efter alder og køn. Stikprøvepopulation i Valgundersøgelserne G@F i forhold til den samlede befolkning 01.01.08.

	Valgundersøgelse G@F (i %)			Hele befolkningen (i %)*		
	Mænd	Kvinder	I alt	Mænd	Kvinder	I alt
18-24 år	4,9	6,7	5,7	10,1	9,4	9,7
25-39 år	18,9	20,8	19,8	27,4	25,8	26,6
40-49 år	16,6	15,7	16,2	19,2	18,0	18,6
50-59 år	20,1	19,4	19,8	17,9	17,1	17,5
60-69 år	18,3	17,8	18,1	14,0	13,9	13,9
70+ år	21,2	19,6	20,4	11,3	15,9	13,6
I alt	100	100	100	100	100	100
N			2.576			nnnn

Note: I datamaterialet er unge underrepræsenterede i forhold til befolkningen som helhed, mens ældre og især mænd over 70 er overrepræsenterede.

*Andel af den danske befolkning 18+, Danmarks Statistik

Tabel 2.3 Køn*alder fordelingen i internetpanelet Gallup@Forum

	Hele G@F (i %)		
	Mænd	Kvinder	I alt
18-24 år	5,6	9,6	7,7
25-39 år	29,0	38,5	34,2
40-49 år	22,5	23,5	23,0
50-59 år	22,1	18,3	20,0
60-69 år	16,4	8,6	12,2
70+ år	4,5	1,6	2,9
I alt	100	100	100
N			49.846

Note: Udtrækket er lavet over samtlige respondenter i TNS Gallups internetpanel Gallup@Forum d. 29/5-2008

Tabel 2.4 Svarfordelingen efter uddannelse. Stikprøvepopulation i Valgundersøgelserne G@F i forhold til den samlede befolkning 01.01.08

	Valgundersøgelse G@F (i %)	Hele befolkningen* (i %)	Hele G@F
Folkeskole/Mellemskole/Realeksamen /præliminær	29,3	33,5	11,6
Studentereksamen/HG/HH/HF/HTX- Eksamen	7,3	8,6	22,4
Erhvervseksamen	34,8	34,4	33,7
Videregående uddannelse	28,5	23,5	43,8
I alt	100	100	100
N	2.576	2.576	49.688

*Andel af den danske befolkning 18+, Danmarks Statistik

Tabel 2.5 Svarfordelingen efter Politisk interesse. Stikprøven i valgundersøgelsen G@F sammenlignet med valgundersøgelserne 2005 samt Gallup-omnibusundersøgelse primo 2002. Andele i pct.

	2002- Omnibus**	2005*	2005****	2008*	2008***
Meget interesseret	19	22,8	21,4	30,4	27,0
Noget interesseret	47	49,2	49,4	50,1	50,5
Kun lidt interesseret	30	24,7	25,8	18,1	20,7
Slet ikke interesseret	4	3,2	3,5	1,3	1,8
I alt	100	100	100	100	100

* Tallene er uvejede.

** Tallene stammer fra internt arbejdspapir af Mogens Storgaard Jakobsen, Gallup A/S, marts 2002, afrundede pct.

*** Tallene er vejede efter køn* alder, region, uddannelse og partivalg 2007.

**** Tallene er vejede ved hjælp af vægten: VEJFV05, som er en vægt der findes i datamaterialet leveret af TNS Gallup. Vægten er generere ved iterativ vejning ud fra variablene køn* alder, geografi og partivalg 2005.

Note: Kategorien "ved ikke" er ej medtaget. Dette ændrer dog ikke ved tallene, da kun to respondenter falder i denne.

Der er en større andel af meget interesserede i forhold til både omnibus 2002 og valgundersøgelsen 2005. Samtidigt ses, at andelen der kun er lidt eller slet ikke interesseret er faldet i forhold til valgundersøgelsen 2005.

TNS Gallup anvender en såkaldt iterativ vejprocedure for at korrigere for de skævheder, der måtte være i data. Det vil sige, at data først vejes så variabel 1 passer med den ønskede fordeling, dernæst vejes der på variabel 2 indtil den har den ønskede fordeling. Dette resulterer i ændringer i fordelingen på variabel 1. Derfor begynder man forfra og vejer på variabel 1 og dernæst variabel 2. Dette gøres indtil vægtene konvergerer, og begge fordelinger passer simultant.

I denne undersøgelse er der anvendt iterativ vejning. Væjekriterierne har været alder*køn², region og uddannelse. I datafilen er der inkluderet fire vægte:

1. Køn*alder (Vægt1)
2. Køn*alder og region (Vægt2)
3. Køn*alder, region og uddannelse (Vægt3)
4. Køn*alder, region, uddannelse og partivalg (Vægt)

² Der er anvendt aldersinddelingerne 18-24 år, 25-39 år, 40-49 år, 50-59 år, 60-69 år, 70+ år.

7 Metodeforskelle og holdningsspørgsmål

De danske valgundersøgelser gennemføres traditionelt som personlige interviews (CAPI). Da indeværende undersøgelse er gennemført som en webbaseret undersøgelse (CAWI) fra et internetpanel er der en række emner, som bør berøres.

For det første ser man ofte, at flere svarer i de 'ikke meningsbærende' svarkategorier som 'ved ikke' og 'ønsker ikke at svare' ved webundersøgelser, fordi denne mulighed er synlig for respondenterne (disse læses som standard ikke op i CATI og CAPI undersøgelser). For at minimere denne metodeforskel var de ikke meningsbærende kategorier skjult for respondenterne i denne undersøgelse. Det var først i det øjeblik, at respondenterne forsøgte at gå videre i spørgeskemaet uden at have besvaret spørgsmålet, at muligheden blev synlig. Resultatet af dette har været en meget lav andel, som har svaret ved ikke. Faktisk er andelen betydelig lavere i denne undersøgelse end ved CAPI undersøgelsen fra 2005 valget. Andelen af 'Ved ikke' svar på fire udvalgte spørgsmål fra nypolitikindekset (Indvandring, miljø, vold og ulandsbistand) ligger mellem 0,1-0,2 % i denne undersøgelse mod 1,8-4,8 % i 2005 undersøgelsen³.

For det andet er skemaet relativt langt at besvare, ca. 55 minutter. Sammenlignet med de undersøgelser, som normalt foretages på internetpanelet, er dette et meget langt skema. For ikke at risikere en meget lav besvarelsesprocent blev der udlovet i alt for 30.000 kr. incentives, hvilket vurderes at have haft en meget positiv effekt på besvarelsesprocenten. Det kan betyde at en særlig gruppe lader sig påvirke af de relative høje incentives⁴.

For det fjerde kan der også opstå uoverensstemmelser mellem CAPI og CAWI undersøgelser i forbindelse med såkaldte grid-spørgsmål. Det vil sige spørgsmålsbatterier, som har samme svarskala og i CAWI undersøgelse vises på samme side. Forskellene består i, at respondenterne i CAWI undersøgelser er betydeligt mere tilbøjelige til, at svare således at korrelationerne mellem spørgsmålene bliver højere. Da der findes relativt mange af denne type spørgsmål i denne undersøgelse kan det udgøre et reelt problem i forhold til sammenligneligheden med tidligere undersøgelser. For at undersøge denne problemstilling sammenlignes korrelationerne mellem udvalgte spørgsmål i Valgundersøgelsen 2005 med de foreliggende data. Princippet er her, at væsentlige forskelle i korrelationerne tjener som indikation på, at der er forskelle mellem de to undersøgelser⁵. Om det er respondenterne eller metoden, der forårsager eventuelle forskelle er dog ikke umiddelbart muligt at afgøre på baggrund af foreliggende data.

³ Der er tale om vejret data i begge undersøgelserne.

⁴ Få dage inde i indsamlingsperioden blev incentives strukturen ændret fra mange små incentives til færre store. Årsdagen var en meget lav svarprocent. Ændringen i strukturen vurderes at have haft en meget positiv effekt på svarprocenten. Incentivestrukturen blev oplyst til alle der klikkede sig ind på linket efter ændringen var foretaget. De som allerede havde besvaret skemaet deltog i samme lodtrækning i den nye incentivesstruktur.

⁵ Analysen er foretaget ved Pearsons R og Kendals tau_b. Man kan ikke sammenligne den numeriske størrelse mellem koefficienterne, da eventuelle forskelle kan være forårsaget af en lang række andre ting end selve indsamlingsmetoden eller respondentsammensætningen (se fx. King 1986, 1989; Achen 1982). Dog er det vel rimeligt i dette tilfælde at argumentere for, at forskellene (fx faktor 2) i korrelationerne mellem spørgsmålene kan tjene som en indikation for at metoden kan tilskrives en effekt.

For at minimere effekten af indsamlingstidspunktet, vælges variable som forventes at have en stabil relation til hinanden. Variablene, som udgør det nypolitiske indeks har vist sig at have en temmelig stabil relation til hinanden over de seneste valg. Disse udvælges derfor til analysen. De enkelte items i nypolitik indekset handler om indvandring, ulandsbistand, straf og miljø.

Tabel 3.1 *Pearson's korrelationer mellem nypolitik items i 2005 og Gallup@Forum undersøgelsen 2008* N 2005 = 2.264. N 2007 = 2.576.

	Indvandring		Miljø		Vold	
	2005	2008	2005	2008	2005	2008
Miljø	0,316	0,369				
Vold	0,392	0,403	0,297	0,276		
Ulandsbistand	0,466	0,583	0,275	0,413	0,349	0,431

Note: Data er vægtet efter samme principper som i de ovenstående tabeller.

Tabel 3.2 *Kendals tau_b korrelationer mellem nypolitik items i 2005 og Gallup@Forum undersøgelsen 2008* N 2005 = 2.264. N 2008 = 2.576.

	Indvandring		Miljø		Vold	
	2005	2008	2005	2008	2005	2008
Miljø	0,264	0,310				
Vold	0,342	0,341	0,226	0,228		
Ulandsbistand	0,420	0,511	0,264	0,361	0,333	0,373

Note: Data er vægtet efter samme principper som i de ovenstående tabeller.

Analysen viser at der er forskelle i korrelationskoefficienterne mellem de to undersøgelser. Indeværende CAWI undersøgelse giver generelt højere korrelationer. Forskellene vurderes dog ikke at være så store at det giver anledning til store bekymringer eller anledning til at kunne konstatere væsentlige forskelle mellem metoderne.

Forskellene i korrelationskoefficienterne kan også være et resultat af ændringer i politiske forhold siden sidste valg, som har givet emnerne/spørgsmålene en lidt anderledes betydning og som derfor influerer på besvarelser og korrelationer. Ud fra foreliggende data er dette ikke muligt at afgøre om forskellene er et resultat af metodeforskelle eller politiske forandringer.

En anden forskel, som man ofte finder mellem metoderne er, at respondenter på CAWI afgiver mere ekstreme svar end på CAPI. En gennemgang af de tre nypolitik spørgsmål viser ikke en tendens til mere ekstreme holdninger ved G@F.

Formålet med denne analyse var ikke at afgøre om CAPI eller CAWI metoden er den bedste, men at give en indikation om, hvilke forskelle der opstår som resultat af om man anvender den ene eller anden metode. Når man sammenligner indeværende undersøgelse med Valgundersøgelsen 2005 findes, der nogle forskelle. Analyserne i dette notat har ikke kunnet give nogen entydig indikation af hvad disse forskelle skyldes. Forskellene kan skyldes faktorer som forandringer i danskernes holdninger, statistiske tilfældigheder og indsamlingsmetode (både udvælgelse og besvarelsesmetode).

8 Litteratur

Achen, Christopher. [*Interpreting and using regression*](#). Fage Publications; Sage University papers, 1982.

Gary King. [*Unifying Political Methodology: The Likelihood Theory of Statistical Inference*](#). Cambridge, England and New York: Cambridge University Press, 1989.

King, Gary. "[How Not to Lie With Statistics: Avoiding Common Mistakes in Quantitative Political Science](#)," *American Journal of Political Science*, Vol. 30, No. 3 (August, 1986): Pp. 666-687.