
제6회 유럽조사연구학회(ESRA) 컨퍼런스 참가 결과보고

6th Conference of European Survey Research Association

2015. 7.

목 차

I. 출장개요	2
II. 주요내용	4
III. 세부내용	6
1. 조사표설계	6
2. 자료연계	12
3. 조사원효과	15
4. 혼합모드	17
IV. 시사점	19
붙임1: 발표자료	21
붙임2: 컨퍼런스 전경	25

I

출 장 개 요

□ 참가 회의

○ 회의명

- 제6회 유럽조사연구학회 컨퍼런스(6th Conference of the ESRA*)

* European Survey Research Association : 조사자료 품질 향상을 위해 유럽 지역과 다른 지역 조사방법 연구자 간 커뮤니케이션을 목적으로 설립되었으며, 격년으로 컨퍼런스 개최

* ESRA 컨퍼런스 사이트 : www.europeansurveyresearch.org/conference

○ 회의 기간 및 장소

- 기 간 : 2015. 7. 13(월)~7. 17(금) (*참가기간: 14~16)
- 개최지 : 아이슬란드 레이캬비크

○ 참가 규모

- 190여개 세션에서 725개 주제 및 60개 포스터 발표
- 유럽, 미국 등 45개국에서 조사방법론 관련 전문가 750여명 참석
. 이 중 아시아권에서는 8개국 20여명 참가(일본, 대만, 중국 등)

□ 출장 수행 내역

○ 통계개발원 수행 연구논문(2건) 발표를 통한 연구성과 확산(붙임1 참조)

- 제목 : The Analysis of Respondent's Behavior toward Edit Messages in Web Survey(박영실 주무관, 7.16.)
. (연구목적) 웹서베이에서 내검 메시지에 대한 응답자 행동 유형 분석을 통해 자료품질향상을 위한 제언

- . (연구내용) 웹 서베이에서의 내검 메시지 발생 현황, 내검 메시지에 대한 응답자 행동 유형 분류 및 내검 메시지 발생에 영향을 미치는 요인을 분석. 응답자들은 내검 메시지 수신 이후 응답내용을 정확하게 수정하는 것 이외에도 추후 메시지 수신을 피하기 위해 여러 전략적 행동을 하는 것으로 나타났다.
- 제목 : Record Linkage Method for the Social Survey in Statistics Korea(정미옥 주무관, 7.15.)
- . (연구목적) 사회변화에 따른 심층적이고 복합적인 통계생산 요구에 부응하기 위해 자료연계 방법 제안
- . (연구내용) 격년으로 주제를 달리하여 실시하는 사회조사 자료를 식별정보를 이용한 결정적 연계와 특성정보를 이용한 확률적 연계를 통해 연계자료를 구축하여 자료의 분석 및 활용가치 제고
- 선진 연구동향 파악을 통한 연구 역량 강화
 - 출장자 연구분야인 자료연계와 조사표설계를 중심으로 혼합모드, 조사원오차, 파라미터 등 다양한 조사방법론 분야의 최신 동향 파악
- 조사방법론 관련 인적 네트워크 구축
 - 국가통계기관 외에 학계 및 민간리서치 회사의 연구자들이 대거 참가하여 연구토론의 장이 활발하게 마련됨에 따라 이들과의 네트워크 구축
 - . Mick Couper(미시건대), Don Dillman(워싱턴대), Dave Tuttle(미센서스국) 등 조사방법론 학자들과 우리나라 연구결과 공유 및 의견 교환

□ 출장자

- (동향분석실) 박영실 주무관, (조사연구실) 정미옥 주무관

II

주요 내용

- 무응답, 가중값 등 전통적인 조사방법 영역은 물론, 웹.모바일 기반 조사방법, 혼합조사, 자료연계, 조사표설계, 조사원오차, 파라미터 등 조사방법론과 관련한 전 영역에 걸쳐 많은 논문들이 발표됨

컨퍼런스 주요 주제

- Design of Response Scales in (Mobile) Web Surveys
- Experimental designs in online survey research
- Using Paradata to Improve Survey Data Quality
- Weighting issues in complex cross-sectional and longitudinal surveys
- Measurement errors in official statistical surveys
- The impact of questionnaire design on measurements in surveys
- Advanced survey estimation methods for treatment of non-sampling errors
- Modelling unit nonresponse and attrition processes
- Poster session
- Analysis of Cognitive Interview Data
- Web and mixed mode data collection in National Statistics
- Linking survey data and auxiliary data sources: statistical aspects and substantive applications
- Longitudinal surveys – challenges in running panel studies
- Recent developments in the analysis of panel data
- Unit and item nonresponse

- 조사표 설계 분야에서는 조사항목 설계가 측정오차에 미치는 효과, 웹 및 모바일 조사표 설계, 인지면접 자료분석, 조사질문의 번역 등에 대한 주제들이 논의되었음
 - 기술발전에 따라 조사항목에서 질적인 자료(개방형)를 얻고자 하는 시도가 있다는 점이 눈에 띈. 동시에 분석과정에서도 질적인

자료에 기반한 심층분석이 많이 이루어지고 있었음

- 자료연계(record linkage)를 통한 조사자료의 품질향상 및 추정값의 정확성 제고, 2차자료 작성, 응답자의 행정자료 활용 동의 방안, 무응답 특성, 기술적 문제 등에 대한 광범위한 논의가 있었음
 - 주로 조사자료와 행정자료 간 정확연계를 기반으로 자료연계 결과를 활용한 응용연구가 많이 진행되고 있음
- 조사원 분야에서는 조사원이 자료에 미치는 영향, 조사과정에서의 조사원 허위행동, 현장조사 과정에서의 조사원 전략 등에 대해 다루었음
- 응답률 향상.비용 절감.조사품질 유지 등 자료수집 방법을 효율화하기 위하여 국가통계에서의 웹과 혼합모드 자료수집 방법, 혼합조사 설계와 모드효과의 추정, 모바일과 웹조사 방법 관련 연구들이 활발히 진행되고 있음

III 세 부 내 용

□ 조사표 설계

◆ 조사항목 설계가 자료 품질에 미치는 효과 관련 연구

(1) Rating scale labeling in web surveys? (독일 사회과학연구소 GESIS)

- (연구배경) 인지절약이론(satisficing theory)에 따르면, 동의한다-동의하지 않는다(agree-disagree)와 같은 척도 문항은 여타 문항에 비해 응답하기 수월하여 조사표 설계 시 많이 활용되고 있으나, 여러 유형의 측정 오차 가능성이 있음. 이 연구에서는 척도값 설명 방법으로 숫자 척도와 언어 척도에 대한 응답자 인지적 부담과 신뢰도에 대해 비교
- (연구가설) 언어 척도가 숫자 척도에 비해 (1) 응답자에게 인지적 부담을 덜 주며, (2) 응답한 내용에 대한 신뢰도의 경우에도 신뢰도가 높을 것이다.

언어척도

does not apply at all	does not apply	applies to some extent	applies partly	applies to a great extent	applies almost fully	applies fully
☐	☐	☐	☐	☐	☐	☐

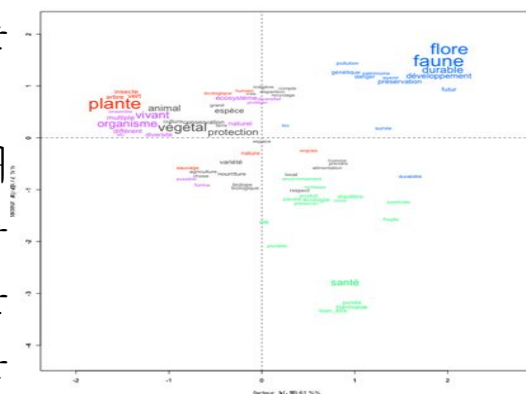
숫자척도

does not apply at all						applies fully
1	2	3	4	5	6	7
☐	☐	☐	☐	☐	☐	☐

- (연구방법) (가설1) 82명 참가자를 대상으로 2012년 Tobii 시선추적 장치를 활용하여 응시시간 및 응시횟수를 분석 (가설2) 2014년 실험 온라인 패널 497명 대상으로 조사 실시 후 주요인분석에 따라 척도 항목 합산 한 후 신뢰도 점수 비교
- (연구결과) 숫자척도가 언어척도에 비해 각 척도를 인지하는 시간은 짧게 나타나나 응답하는 시간은 더 길게 나타났음. 주요 항목에 대한 신뢰도 계수에서도 언어척도 항목이 그 점수가 더 높아 두 가설을 모두 지지하는 것으로 나타났음

(2) Exploring a new way to avoid errors in attitude measurements due to complexity of scientific terms (스위스 로잔대학교)

- (연구배경) 과학적 전문 용어를 과연 일반인들이 어떻게 처리할 것인가?에 대한 문제제기에서 출발하여, ‘어려운 개념’에 대한 이해과정에서 발생할 수 있는 측정오차를 제어할 수 있는 방법론적 접근법 제안
- (연구가설) (1) 과학 용어 중 ‘biodiversity’에 대한 개인들의 태도를 측정하기 위해 사회적으로 대표성(social representation) 있는 용어를 생산해 낼 수 있는지?, (2) 이것의 신뢰도와 타당성이 어떻게 평가될 수 있는가?
- (연구방법) 제네바 거주자 중 357명을 추출하여 biodiversity라는 용어를 처음 들었을 때 떠오른 표현이 무엇인지를 종이와 웹조사를 통해 개방형으로 질문한 후 그 결과를 텍스트 분석프로그램으로 분석하여, 대표성 있는 질문으로 재구성하여 2차 조사(512명-종이)를 통해 타당성 검증
- (연구결과) (가설1) biodiversity에 대해 오른쪽 그림과 같이 5가지 의미 그룹(색깔로 구분)이 발생하였으며, 이것을 이용하여 개념정의 후 5점 척도 질문으로 재구성하였음. (가설2) 폐쇄형



질문으로 재구성한 버전1과 개방형 질문으로 되어 있는 버전2에 대해 분석한 결과 가설1에서 보여준 개념과 버전2에서 보여준 개념구조가 유사하였음(가설2-1). 그런데, 버전1과 달리 개념정의가 되어 있지 않은 버전2에서 덜 교육 받은 사람들의 응답이 중간척도에 몰리는 경향이 있었음. 또한, 버전 1의 태도 척도가 행동 척도를 예측하는 추가적인 정보를 보유한 것으로 나타나 전반적으로 자료의 품질에 긍정적인 효과가 있는 것으로 분석되었음

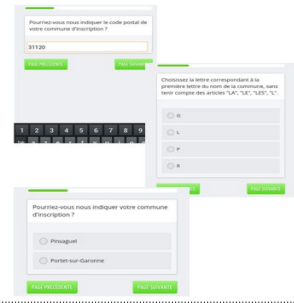
(3) Measuring attitudes towards immigration with direct questions –
can we compare 4 answer categories with dichotomous responses?
(에스토니아 탈린대학교)

- (연구질문) 민감한 질문을 물을 경우, 질문을 묻는 방식과 사회적
인 바람직성이 응답분포에 강하게 영향을 끼치는지? 그리고 그것의
영향이 교육이나 연령에 따라 달라지는지?
- (연구방법) 이민에 대한 태도 관련 3가지 질문에 대해 응답범주를
조건 1에서는 이분형(허가 vs 비허가) 조건2에서는 4점 척도로 허가
여부를 물었음
- (연구결과) 이민을 반대하는 경향은 응답척도에 강하게 영향을
받았는데 4점척도에 비해 이분형 척도에서 30%에서 50%까지 비허가
입장을 보여주었음. 교육수준이 낮은 경우 중간범주를 선택할
가능성 또한 낮았음

◆ 웹 및 모바일 조사표 설계

(1) Adapting survey questionnaires to touch-screen tablets (프랑스 국제비교연구대학)

- (조사설명) ELIPSS(Etude Longitudinal par Internet Pour les Sciences Sociales) 패널의 모바일 조사표 구성 → 프랑스에 거주하는 18-75세 1000명의 확률표본, 응답자기입식으로 인간 및 사회과학 주제의 조사질문으로 구성 (30분가량 소요)
- (표준 조사표안 구성)

페이지 스타일	터치 스크린	제한적 스크롤링	드롭다운 메뉴
			
자동완성기능	질문분할	응답방법 설명	주관식
			

- (주요 실험 결과)

▶매트릭스형 질문 vs 나열식 질문

매트릭스형	페이지별 나열형																								
Et au cours de votre vie, avez-vous déjà... ? <table border="1"> <thead> <tr> <th></th> <th>Oui</th> <th>Non</th> </tr> </thead> <tbody> <tr> <td>Utilisé un moteur de recherche (Yahoo, Google, Bing, etc.)</td> <td>☐</td> <td>☐</td> </tr> <tr> <td>Envoyé des e-mails avec des fichiers joints (document, photo)</td> <td>☐</td> <td>☐</td> </tr> <tr> <td>Posté des messages dans un forum de discussion en ligne ou lors d'un «chat»</td> <td>☐</td> <td>☐</td> </tr> <tr> <td>Téléphoné en vous connectant à Internet (Skype, MSN Messenger, Hangout, etc.)</td> <td>☐</td> <td>☐</td> </tr> <tr> <td>Créé une page web</td> <td>☐</td> <td>☐</td> </tr> <tr> <td>Mis en ligne des textes, des jeux, des photos (sur des réseaux sociaux par exemple)</td> <td>☐</td> <td>☐</td> </tr> <tr> <td>Modifié les règles de sécurité de votre navigateur Internet</td> <td>☐</td> <td>☐</td> </tr> </tbody> </table> PAGE PRÉCÉDENTE PAGE SUIVANTE		Oui	Non	Utilisé un moteur de recherche (Yahoo, Google, Bing, etc.)	☐	☐	Envoyé des e-mails avec des fichiers joints (document, photo)	☐	☐	Posté des messages dans un forum de discussion en ligne ou lors d'un «chat»	☐	☐	Téléphoné en vous connectant à Internet (Skype, MSN Messenger, Hangout, etc.)	☐	☐	Créé une page web	☐	☐	Mis en ligne des textes, des jeux, des photos (sur des réseaux sociaux par exemple)	☐	☐	Modifié les règles de sécurité de votre navigateur Internet	☐	☐	Et au cours de votre vie, avez-vous déjà : - Utilisé un moteur de recherche (Yahoo, Google, Bing, etc.) ? ☐ Oui ☐ Non PAGE PRÉCÉDENTE PAGE SUIVANTE
	Oui	Non																							
Utilisé un moteur de recherche (Yahoo, Google, Bing, etc.)	☐	☐																							
Envoyé des e-mails avec des fichiers joints (document, photo)	☐	☐																							
Posté des messages dans un forum de discussion en ligne ou lors d'un «chat»	☐	☐																							
Téléphoné en vous connectant à Internet (Skype, MSN Messenger, Hangout, etc.)	☐	☐																							
Créé une page web	☐	☐																							
Mis en ligne des textes, des jeux, des photos (sur des réseaux sociaux par exemple)	☐	☐																							
Modifié les règles de sécurité de votre navigateur Internet	☐	☐																							

- (결과) 무응답자 비율은 매트릭스형이 6.0%로 나열형보다 높고, 무응답 항목은 매트릭스형이 4.5개로 나열형 1.1개 보다 높음

- (개선) 매트릭스 보다는 질문을

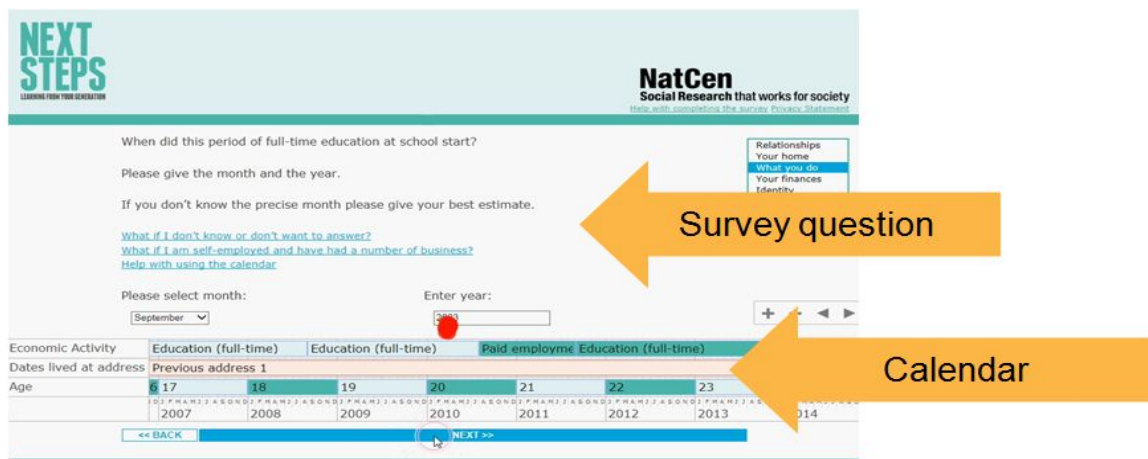
여러 개로 나누어 제시
 ▶DK(모르겠다) 범주 제시 여부

비제시	제시
	

- (결과) 무응답률이 DK 범주를 제시한 경우 11.3%로 제시하지 않은 경우 1.5% 보다 높았음
- (개선) 중요 질문에 대해서는 DK 범주를 제시하지 않았으며, 민감한 질문이나 사실적 질문에 대해서만 제시

(2) Developing a web-administered Event History Calendar (영국 종단조사연구소)

- (연구배경) 종단조사에서 생애사건달력은 회상오류 감소, 자료 간 논리적 일관성 증진, 응답 도움 역할을 하는데 기여



- (연구방법) 주소, 동거여부, 경제활동상태 등에서 변화를 겪은 23-27세 11명에 대한 시선추적실험 및 인지면접 수행
- (연구결과) (가설1) 오직 일부의 참가자가 생애사건달력에 대한 지시문을 대강 보는 경향이 나타났으며, (가설2) 실제 응답하는 과정에서도 해당 달력을 전혀 보지 않는 경우와 간헐적으로만 보는 집단으로 구분되었고, (가설3) 참가자들은 생애사건달력 없이도 정확하고 완전한 정보를 응답할 수 있다고 생각하는 경향이 있었음. 그러나, 응답내용에서 오류 발견 → 향후 생애사건달력의 기능에 대한 비디오 설명을 추가하고, 또한 사용방법 보다는 왜 도움을 줄

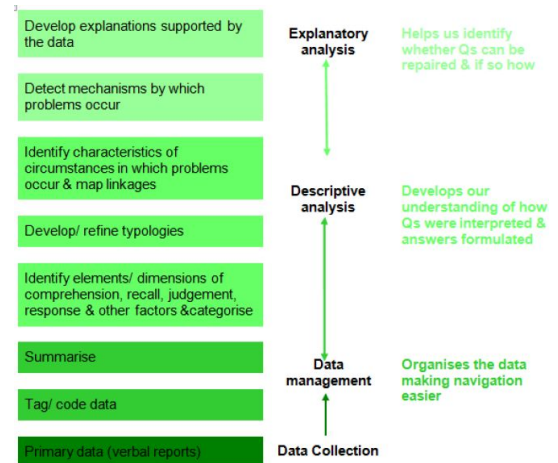
수 있는지를 설명할 것을 권고

◆ 인지면접 자료 분석

(1) I don't believe it : How credible are your cognitive interview findings? (영국 사회연구소)

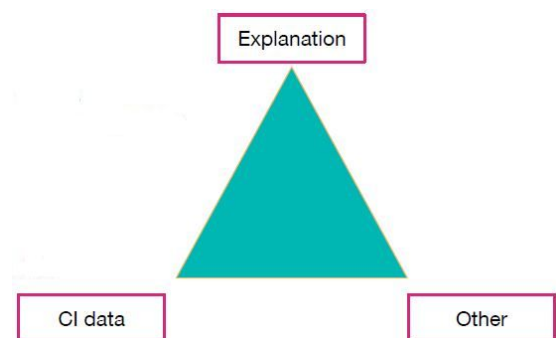
- (연구배경) 인지면접은 소규모, 자발적 지원자를 대상으로 프로빙 등의 기법을 통해 얻어낸 사고과정을 통해 숨겨진 문제를 식별한다는 특징을 갖고 있으나, 모든 사람이 자신의 사고과정에 대해 발화할 수 없고, 식별된 문제의 규모를 말하기 어렵다는 한계

- (인지면접 자료 분석과정) 자료수집 → 자료관리→기술적 분석→설명적 분석



<그림> 인지면접 자료 분석과정

- (주요 질문) 1) 인공적 결과물(artefacts)에 대해 주의하라: 언제 구술 보고서가 작성되었고, 어떠한 기법을 사용하여 문제를 도출하였으며 사전에 논의된 적은 없는지, 이 결과가 서베이 세팅에서도 반복될 수 있는지? 2) 도출된 결과가 여타의 조사표 관련 이론 혹은 자료와 정합되는가? 3) 1회성 결과물은 어떻게 다루었는가: 혹시 너무 적게 면접을 했거나 하는 등의 샘플링 이슈와 관계는 것은 아닌지? 큰 그림을 그리는데 포커스



<그림> 인지면접자료와 이론과의 관계

- (연구결과) 설명적 분석과정에서는 인지면접이 갖고 있는 한계를 인식, 인위적 효과 및 설명가능한 추가 자료 및 자료품질에 미치는 효과에 대해 인지할 것

□ 자료연계

(1) Challenges and opportunities of making use of income registers in the EU-SILC weighting procedure (오스트리아 통계청)

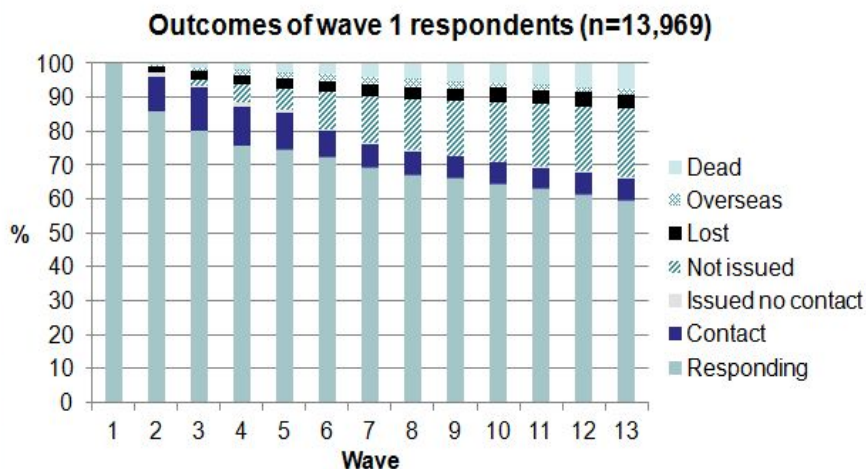
- (연구배경) 오스트리아 EU-SILC(Statistics on Income and Living Conditions) 조사에서는 등록자료와 연계하여 소득 관련 정보를 획득하고 있음
- (연구목적) 연계된 등록자료의 소득정보를 가중값 작성에 활용: 단위 무응답 편향 분석에 활용 가능한지를 검토
- (연구내용) 오스트리아의 EU-SILC조사는 횡단과 종단조사를 목적으로 연동패널로 설계되었고, 표본 가구의 자발적 참여에 의해 수행되고 있음. 최소 유효 표본 수는 4,500가구이고, 단위 무응답이 2014년 76%임. 연방정부로부터 표본 전체에 대한 암호화된 개인 식별자(bPK)를 제공받아 등록자료의 소득 정보를 연계함
 - ① 등록자료(wage tax register)의 소득 정보 주변분포를 이용하여 새로운 조정가중값 작성(칼리브레이션); 임금소득자수, 노령급여소득자수
 - ② 선택적 단위 무응답에 의해 소득지표(AROP)에 편향이 발생하는지 분석. 응답확률만의 모형과 접촉확률과 응답확률의 모형비교 분석
- (연구결과) 2단계 모형(접촉확률과 응답확률의 모형)이 편향을 다소 감소시키는 것으로 나타남. 설계가중값에 직접적으로 적용하더라도 칼리브레이션이 편향을 축소시킬 수 있음

(2) Dead or Alive? Dealing with Unknown Eligibility in a Longitudinal Survey (호주 멜버른대학교)

- (연구배경) 종단조사에서 웨이브가 증가함에 따라 추적실패나 대상 외(out of scope)등의 이유로 접촉불가(non-contact)한 표본이 증가.

그 중 사망으로 인한 경우는 얼마나 차지하고 있으며, 이러한 사망이 응답률과 가중값에는 어떠한 영향을 미치는지 알지 못함

- (연구목적) 호주의 대표적 패널조사인 HILDA(Household, Income and Labour Dynamics)조사에 대해 알려지지 않은 적격(unknown eligibility)에 대해 네 가지 방법을 실험
- (연구내용) 1차 웨이브의 응답자의 25%가 13차 웨이브에서 접촉불가



- ① 방법1: 사망등록자료와 연계하여 사망자 파악. 연계방법은 이름, 출생일, 성별 순으로 가중값을 계산하여 가장 높은 가중값을 갖는 사람을 연계. 사망으로 새로 밝혀진 경우들을 모두 내용검토. 사망으로 알려진 경우도 등록자료와 5%는 연계되지 못함(해외사망, 개명, 출생일 오류)
 - ② 방법2: 생명표(life table)이용하여 나이, 성별, 연도별 사망 예측.
 - ③ 방법3: 조사자료를 이용하여 나이, 장기간 건강상태, 혼인여부, 고용상태에 따른 생존곡선 모형 적용
 - ④ 방법4: 1차부터 13차까지 응답한 표본을 분류하여 가중값 조정
- (연구결과) 등록자료를 연계한 방법이 5%의 손실이 있지만 가장 정확. 단기간으로는 방법3이 우수, 장기간으로 방법2가 우수

(3) Using administrative data about employment status to improve the Austrian LFS estimation (오스트리아 통계청)

- (연구배경) 과거 LFS에서는 지역별 성*연령, 국적, 가구원수의 주변 분포를 이용하여 가구 내 가구원은 모두 동일한 값을 갖도록 가중값을 작성해왔음. 2011년 등록센서스를 실시하면서 새로운 인구와 주택 구조를 반영하고, 선택적 무응답에 의한 편향 조정이 필요
- (연구목적) 고용상태 관련 등록자료를 이용하여 고용통계 추정값을 개선
- (연구내용) 등록자료의 고용관련 정보를 이용하여 새로운 주변분포 (지역*성*고용상태)를 작성하여 가중값 조정. 2013년 무응답률은 5.7%로, 모집단에 비해 무응답층의 실업과 비경의 비율이 높고, 취업 비율은 상대적으로 낮게 나타나 무응답에 대한 가중값 조정 실시
- (연구결과) 새로운 가중값의 적용을 통해 실업에 대한 편향을 축소

(4) 조사자료, 암 등록자료, 행정자료 간 자료연계 타당성 연구(이탈리아)

- 세 가지 출처*의 자료 간 연계를 통해 사람의 건강과 사회-경제적 조건들에 대한 상세한 자료를 구축
 - * ① 암 등록자료 ② ISTAT의 조사자료 ③ Umbria 지역의 행정자료
- 후속적으로 이탈리아 전 지역에 대해 연구를 확장할 계획

(5) 조사자료, 행정자료, 상업적 주택자료 간 자료연계 연구(미국 센서스국)

- 세 가지 출처*의 자료를 이용하여 주택 단위의 자료연계를 시도
 - * ① American Housing Survey ② 상업적 주택자료 ③ 주택도시개발부의 행정자료
- 주소와 주택 간 매칭률이 지역 간 차이가 있고, 다가구 주택구조에서 낮은 매칭률을 보임
- 주소의 매칭률을 높일 수 있는 방안과 그 장단점에 대해 논의

□ 조사원 효과

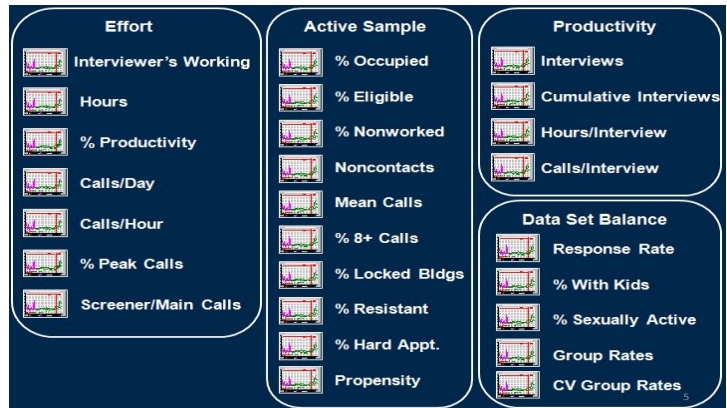
(1) Audio-recording of verbatim thinkalouds: a solution to the problems of interviewer transcription? (영국 사우스햄턴대학교)

- (연구배경) CAPI의 도입은 지금까지 조사원이 응답내용을 적어왔던 (typed) 방법에서 벗어나 응답자의 답변내용을 그대로 녹음하도록 하는 것을 가능케 함. 이는 조사원 부담을 줄여주고 자료 품질을 향상시켜줄 것으로 기대되나, 응답자료가 복잡해지고 응답자 동의가 이루어지지 않는 등의 부정적 효과가 내재할 가능성이 존재
- (연구방법) 과학용어에 대한 개방형 질문을 다루고 있는 2012 Wellcome Trust Motior 조사 대상, 확률표본에 근거한 1396명의 성인이 응답하였으며, 응답자는 무작위로 Audio-recording과 Typed 조건으로 할당, 두 집단 간 세 가지 품질지표(사용된 단어수, 코드화된 자료 수, 단어의 길이)에 차이가 있는지를 비교
- (연구결과) 전반적으로, 응답자들은 과학적 지식의 수준 및 연구 참여 의지 정도에 따라서 자신의 응답내용이 녹음되는 것에 대한 수용 여부에 차이를 보임 (가설1) Typed 집단과 비교할 때, Audio 집단에서 더 많은 단어를 사용하고, 더 많이 코드화되었으며, (가설2) Typed 집단에 비해, Audio 집단에서 조사원에 따른 응답내용의 변동성 또한 감소하였음

(2) Using paradata to monitor interviewer quality (미국 미시건대학교)

- (연구배경) 파라데이터를 활용한 전국 및 조사원 수준에서의 자료 품질 모니터링을 통해 업무 패턴에 대한 이해 및 진행 과정 변경에 대한 빠른 의사결정 가능
- (조사원 모니터링 위한 파라데이터 자료) 표본관리시스템, 조사원 및 조사표로부터의 일일 자료로부터 현장조사시간, 무응답, 도움키

의 사용, 비고입력란의 활용 변수에 대한 요인 분석을 통해 factor1) 얼마나 빨리 진행되는지, factor2) 얼마나 예리가 많은지, factor3) 얼마나 무응답이 많은지 진단



<그림> NSFG 조사의 현장조사 모니터링 자료

- (모니터링과정) 최초에는 소규모 그룹에 대해 검토를 실시 → 문제로 지적된 조사원과 토의 → 조사원 수준에서의 개입여부 판단 → 최종 산출물에 대한 지속적 모니터링

주요 지표별 문제 진단 (적색이 짙을수록 심각)

Average of Zscore	Column Labels						
	C8	Y2	Q01	Q02			
Row Labels	W08	W10	W12	W06	W08	W10	W12
Margolis, Suzanne							
avg_backup_perfield_z	-0.26	-0.35	-0.42	-1.26	-1.41	-1.44	-1.41
avg_DK_perfield_z	1.11	1.37	1.40	0.55	0.77	1.23	1.41
avg_err_esc_perfield_z	-0.38	-0.36	-0.42	-1.12	-1.17	-0.79	-0.79
avg_err_jump_perfield_z	0.93	0.62	0.58	-0.11	0.21	0.70	0.70
avg_err_supp_perfield_z	4.85	4.82	4.86	-0.49	0.18	0.71	0.71
avg_fieldtime_pervisit_z	-0.54	-0.51	-0.49	0.02	-0.06	-0.06	-0.06
avg_qhelp_perfield_z	-0.92	-0.75	-0.77	0.11	0.11	-0.09	-0.09
avg_remdk_perfield_z	2.89	2.91	2.99	2.01	2.58	3.30	3.30
avg_RF_perfield_z	-0.49	-0.42	0.61	-0.30	0.10	0.71	0.71

조사원별 문제 진단

The spreadsheet displays a list of survey workers and their performance metrics across different time periods. The data is organized into columns for worker names and various Z-score categories.

- (문제해결방안) 문제해결 테크닉에 대한 재교육, 모니터링 조사원 숫자 증가, 조사원 그룹전체에 대한 재교육, 개별 면접 케이스 수준에서의 점검

□ 혼합모드

(1) Mixed mode data collection in national statistics: findings from the ESSnet DCSS(네덜란드 통계청)

- (연구배경) 웹 자료수집 도구의 개발과 혼합모드 자료의 효과를 평가 및 구현하기 위해 2012~2014년 까지 유럽 각국이 컨소시엄을 이뤄 노동력조사(LFS)를 사례에 대한 프로젝트를 수행
- (연구목적) DCSS(Data Collection for Social Survey) 프로젝트에서 얻은 경험과 시사점을 공유
- (연구결과) 유럽 사회조사의 절반 이상이 CAWI 구현에 대한 계획이 있음. LFS에 혼합조사를 적용하는 이유로 비용절감, 응답자 편의, 응답률 제고를 위한 것으로 나타남.

Design	LF S1	LF S2+	EU-SILC	ICT	Census	HBS
Single mode						
Capi	10	2	11	4	2	4
Cati	4	6	1	8		
Cawi						1
Papi	8	4	7	6	5	8
Papi + other		1			1	
Pap			1	1	4	3
Registry					8	
Capi + registry	2		1			
Cati + registry	3	3	3	2		
Cati - other + registry						1
Mixed mode interviewer - interviewer						
Capi - Cati	2	8		4		
Cati - Capi - Papi	2	2	2	1		
Cati - Papi		2	1			
Capi - Papi	1	1	1			
Cati - capi + registry		2	4	1	1	
Mixed mode interviewer - noninterviewer						
Capi - pap	1	1				5
Cati - pap - other						1
Cawi - cati - pap				1		
Cawi - capi - papi						1
Cawi - papi					2	
Papi - pap			1	1		6
Cati - capi - papi - pap				1		
Cati - papi - pap						
Capi - papi - pap				1		
Cawi - papi - pap					1	
Cawi - capi - cati - registry	1				1	
Cati - capi - pap - registry						1
Cawi - cati - registry		1				
Papi - Pap - registry - other				1	1	
Cawi - capi - registry					2	
Cawi - Capi - Papi - registry					1	
Cawi - Capi - Cati - pap - registry						1
Mixed mode noninterview - noninterview						
Cawi - pap				1	4	1
Cati - pap						
Papi - pap						
Cawi - pap - registry					1	

최고의 혼합조사 설계는 없으며, 혼합조사 방법은 설계 목적에 따라 결정되며, 아울러 조직, 규제, 문화적 맥락을 반영함. ① 커버리지 오차나 무응답 오차를 줄이는 것이 목적이라면, 조사방법을 선택하게 하는 병행적 설계 또는 비용이 높은 조사방법을 첫 번째 방법으로 하는 순차적 설계를 적용. ② 비용 절감이 목적이라면, 비용이

낮은 조사방법을 첫 번째 방법으로 하는 순차적 설계를 적용.

서로 다른 조사방법에 의해 수집된 자료의 측정오차는 매우 중요하다, LFS의 경우 조사방법(우편, 웹, CATI, CAPI)간 차이는 대부분 공통 가중값 작성 변수에 의해 설명이 됨. 그럼에도 불구하고, 모드 효과는 매우 중요하게 다뤄져야 하고, 측정오차에 대한 조정 (adjustment) 연구는 앞으로 지속적으로 수행되어야함.

(2) Sequential Mixed Mode Experiment in the U.S. National Monitoring the Future Study(미국 미시건대학교)

- (연구배경) MTF(Monitoring The Future)조사에 혼합모드 적용을 위한 시험조사를 실시하고 그 결과를 분석
- (연구목적) 서로 다른 실험조건에 따라 전체 응답률, 특성그룹별 응답률, 모드선택(웹 vs 종이조사), 자료 품질(항목무응답), 추정값, 비용 등에 차이가 있는지 검토
- (연구내용) MTF조사는 전국표본으로 16,000명을 조사하는데, 일부 표본(5,000명)에 대해 웹조사와 우편(종이)조사의 3가지 조건에 대한 순차적 혼합모드 적용하고 결과를 분석. ① MAIL PUSH(우편조사 먼저, 후에 둘 다) ② WEB PUSH(웹조사 먼저, 후에 둘 다) ③ WEB PUSH 및 e-mail로 초대장과 독촉장 전송
- (연구결과) 전체 응답률은 WEB PUSH가 낮았으며, 특성그룹별로는 특권층(백인, 4년제 대학교 희망자, 부모교육정도 대학이상)의 응답률이 낮았음. 모드 선택은 WEB PUSH+ Email 방법인 경우 보다 웹으로 응답하는 경향이 높았음. 조사방법에 따라 결측자료에 차이는 없었음. 음주에 대한 추정값이 WEB PUSH+ Email 방법이 높게 나타남. 비용측면에서도 WEB PUSH+ Email 방법이 가장 효율적임. 이러한 결과에 따라 WEB PUSH+ Email 방법의 선정이 가장 유력하지만, 2016~17년 해당 방법에 대한 연구를 더 수행할 계획임

IV

시 사 점

○ (전반적 인상)

- 이번 유럽조사연구학회에서는 비표본 오차의 영역(조사표설계, 조사원효과, 자료수집방법, 파라데이터 등)에 대한 연구 및 다양한 방법을 활용한 자료수집(행정자료 활용, 자료연계, 모바일 조사 방법)에 대한 연구가 많이 이루어졌음
- 전통적으로 다루어져 오던 확률적.양적 기법과 함께 최근에는 비 확률적.질적 기법 또한 많이 연구되고 있음

□ 전통적 방식에서 벗어나 새로운 도전과 기회의 영역인 통계생산 방법의 선진화에 대한 많은 관심과 연구가 필요

○ (조사표 설계)

- 인지면접 등의 조사표 평가 분야와 함께 조사항목 설계(매트릭스 디자인, 척도문항의 제시방법, 웹조사에서의 텍스트박스 크기 등)가 자료품질에 미치는 효과에 대한 연구가 여러 국가에서 실시되고 있으며 이를 근거로 조사표가 작성되고 있음
- 특히, 웹 및 모바일 조사의 확대에 따라 이에 대한 자료수집방법 적합형 조사표 개발 연구가 한창 진행 중임

□ 2013.5월 구축된 통계개발원 조사표실험실에서는 기존 조사표 개선을 위한 인지면접, 시선추적실험 등의 연구를 중심으로 진행해 오고 있음

□ 향후, 조사항목 설계에 대한 지침을 줄 수 있는 연구 뿐 아니라 조사환경 변화에 대응하기 위한 모바일 조사표 연구도 추진되어야 할 것임

○ (자료 연계)

- 조사자료와 행정자료 간의 정확연계를 기반으로 한 연구들이 추진되면서, 연계된 자료를 이용하여 부가적인 정보를 얻는 것은 물론 조사자료의 품질평가를 통한 품질향상, 추정값의 정확성 제고, 무응답 및 회상오차 등 측정오차 평가 등 조사결과를 개선하기 위한 연구도 매우 활발히 진행 중임

- 또한 표본조사 시 사전에 응답자에게 행정자료 활용에 대한 동의여부를 선택하도록 함으로써 개인정보 이용에 대한 근거 및 연계자료 구축 활성화를 위한 기반을 마련하고 있음

□ 통계개발원에서는 그동안 조사자료와 조사자료 간 정확 연계 및 확률적 연계 방법과 연계된 자료를 이용한 2차자료 작성 연구를 진행해 왔음

□ 올해 등록센서스 도입으로 행정자료 활용이 본격화됨에 따라 조사자료와 행정자료 간 자료연계 연구와 행정자료를 연계하여 조사자료를 개선하는 통계적 방법론 연구 등을 추진해야 할 것임

○ 유럽조사연구학회는 세계 각 국가의 조사방법론 관련 주요 이슈 및 연구 경험을 한 자리에서 공유할 수 있다는 점에서 의미있음

- 아시아 국가 참석률은 여전히 낮은 편이나 2년 전에 비해 일본의 참석이 두드러짐

- 아시아 국가 간 조사방법 연구 결과 공유할 수 있는 장(場)의 필요성이 제기됨

【붙임1】 발표자료

□ 자료1 : The Analysis of Respondent's Behavior toward Edit Messages in Web Survey(박영실 주무관, 7.16.)

The Analysis of Respondent's Behavior toward Edit Messages in Web Survey

Youngshil Park
Statistical Research Institute
Statistics Korea

Contents

1. Introduction
2. Overview of the Internet Census
3. Method
4. Results
5. Discussion

Contents

1. Introduction
2. Overview of the Internet Census
3. Method
4. Results
5. Discussion

The Purpose of Data Editing

From the viewpoint of data quality

To reduce nonsampling errors

- nonresponse errors
- measurement errors

Data Editing in Web Survey

WHEN At the time of data collection	WHERE The place of collecting data	WHO Respondent
WHAT Response Error	HOW Interactive with computer	WHY To improve data quality

Core Questions Regarding Editing

How many edit messages are too many?

It is **difficult to devise empirical rules** to answer since each data collection situation is different (Anderson et al, 2003)

We attempted to **keep a balance** between the **number of edit messages** presented to respondents **and the necessity of** having respondents correct an error (Laroche, 2011)

Core Questions Regarding Editing

How should edit messages be presented to respondents for maximum effect?

Despite the wide use, importance, and potential impact of edits on respondent burden, there has been **little research** on factors that impact the effectiveness of edits.

(Mockovak, 2007)

Summary of Literature Review

General design guidelines based on developer experience and usability studies has been shown rather than empirical studies.

(Anderson et al., 2003)

However,

Do respondents behave as expectation of survey designers?

Satisficing Theory (Krosnick, 1991)

Respondents are likely to **give easy answers without much thought** as they are fatigued and distracted...

Satisficing Behaviors by Krosnick

Weak Satisficing : Selecting the first response alternative that seems reasonable, Agreeing with assertions

Strong Satisficing : Non-differentiation in using rating scales, Saying don't know etc..

Respondents will satisfice or not depending on the relative level of

- Their ability and motivation
- Task difficulty

Research Questions

How many edit messages were displayed?

Was the correct action taken on the first attempt?

What kind of actions do they take regarding edit messages?

Which factors affect to generate edit messages?

Framework of Analysis

```

graph LR
    A[1st message] --> B[2nd message]
    B --> C[3rd message]
    C --> D[...]
    B -- Success --> B1[Success]
    B -- Failure --> C
    C -- Success --> C1[Success]
    C -- Failure --> D
    D -- Success --> D1[Success]
    D -- Failure --> E[...]
    
```

Concept : "success" and "failure"

Success

- Desirable behavior: Rs correct an error
- Satisficing behavior: Rs falsify answers to pass the edit message

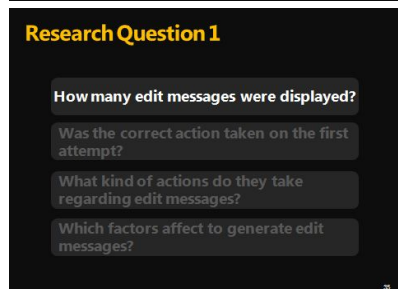
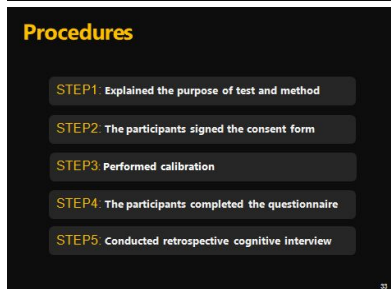
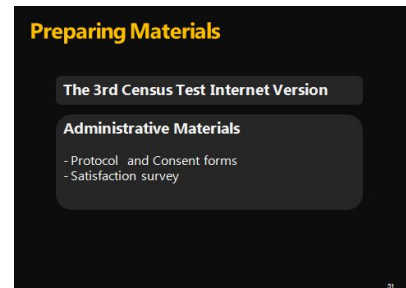
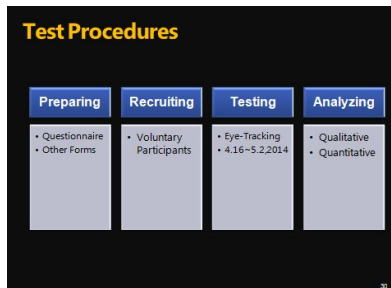
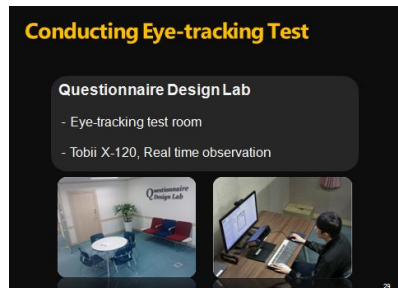
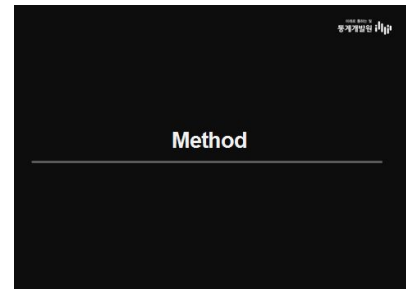
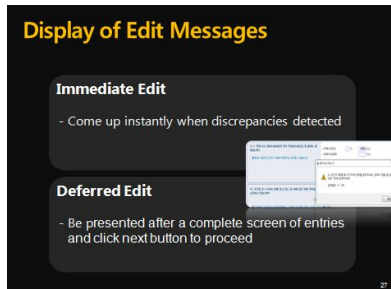
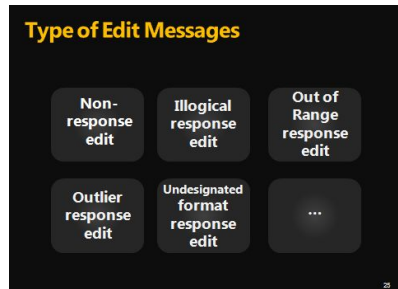
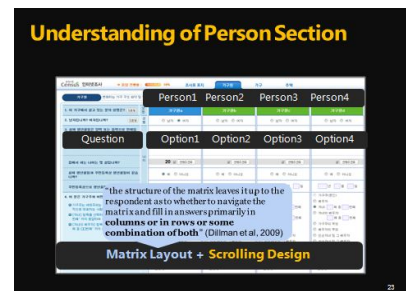
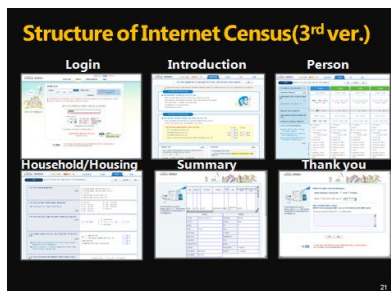
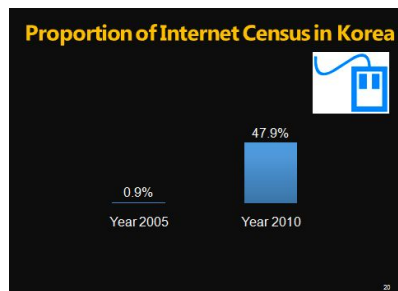
Failure

- Rs don't modify answers despite of edit messages or fail to enter appropriate value

Forms of Satisficing Behavior

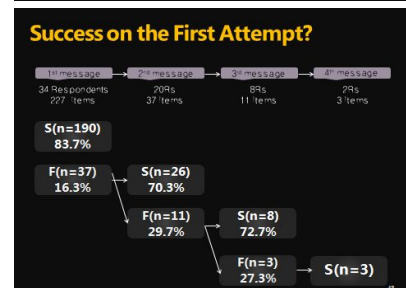
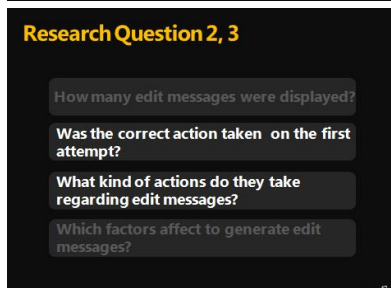
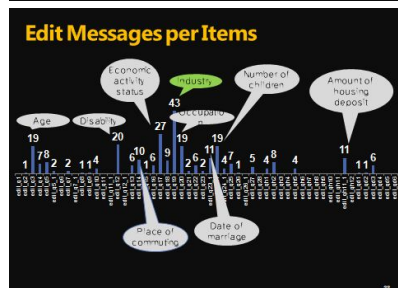
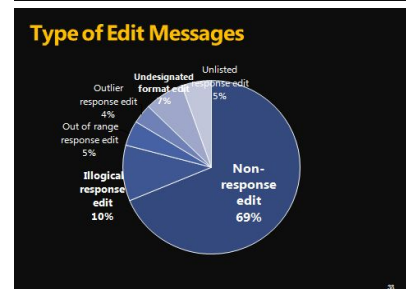
Strong satisficing	Weak satisficing
Reducing number of questions using branching questions	Answering approximately instead exact value
Avoiding open-ended answer in other-specify questions	

Overview of the Internet Census

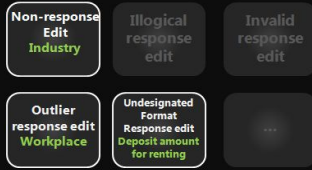


The Number of Edit Messages

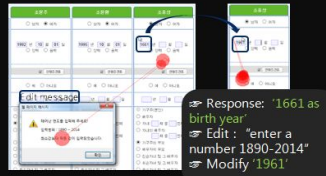
	N	MIN	MAX	M	SUM	SD
Total edit	34	1	24	8.2	278	5.3
Deferred edit	34	0	20	5.4	185	4.5
Immediate edit	34	0	10	2.7	93	2.5



Remained Items on the Final Editing



Success : Desirable Behaviors



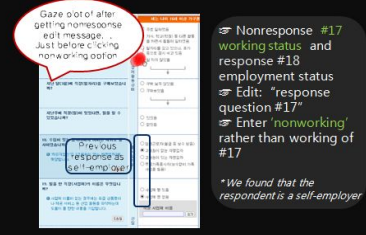
Unintentional Positive Effect

Learning effect

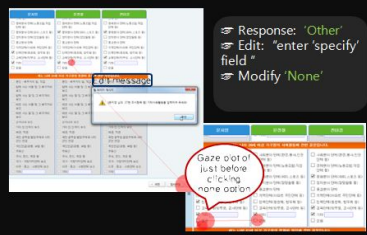
Enter following near nonresponse item with one nonresponse edit message



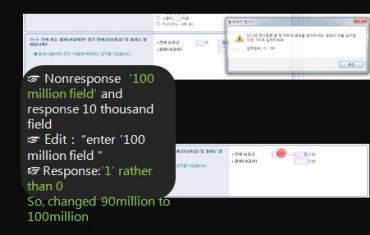
Success : Strong Satisficing Behaviors



Success : Strong Satisficing Behaviors



Success : Weak Satisficing Behaviors



Failure Behaviors

Don't know what the problem is

Ex) error from consistency check between two questions

Don't know how to fix it

Ex) numerical entries

Difficulty of finding the item for editing

Ex) similar topic and structure questions in a sequence



Research Question 4

How many edits messages were displayed?

Was the correct action taken on the first attempt?

What kind of actions do they take regarding edit messages?

Which factors affect to generate edit messages?

Measuring Variables

Task Difficulty

Response pattern, Number of household member

Respondents' Abilities

Age, Education, Hours of computer use per day

* Motivation is assumed to be consistent in voluntary participants

Task Difficulty & Number of Edit

		N	Total		Deferred	
			M	SD	M	SD
Number of household member	1	3	3.3	2.3	2.7	2.1
	2	10	6.5	3.7	3.9	3.6
	3	5	10.8	6.1	7.6	4.3
	4	12	8.3	6.0	5.8	5.2
	6	4	12.3	3.9	7.5	5.6
Response Patterns	vertical	8	5.3	4.1	2.8	2.5
	horizontal	14	8.4	6.0	5.9	5.0
	mixed	12	9.8	4.7	6.7	4.6
Total		34	8.2	5.3	5.4	4.5

Task Difficulty & Number of Edit

		N	Total		Deferred	
			M	SD	M	SD
Number of household member	1	3	3.3	2.3	2.7	2.1
	2	10	6.5	3.7	3.9	3.6
	3	5	10.8	6.1	7.6	4.3
	4	12	8.3	6.0	5.8	5.2
	6	4	12.3	3.9	7.5	5.6
Response Patterns	vertical	8	5.3	4.1	2.8	2.5
	horizontal	14	8.4	6.0	5.9	5.0
	mixed	12	9.8	4.7	6.7	4.6
Total		34	8.2	5.3	5.4	4.5

Respondents' Abilities & Number of Edit

		N	Total		Deferred	
			M	SD	M	SD
Age	10-20's	24	6.9	5.2	4.0	4.1
	40-50's	10	11.3	4.3	8.9	3.8
	university+	7	9.9	5.5	7.9	5.4
Education	university+	27	7.7	5.3	4.8	4.2
	1hour+	14	9.7	6.2	7.1	5.7
Computer use per day	1 hour+	20	7.1	4.4	4.3	3.2
	1 hour+	20	7.1	4.4	4.3	3.2
Total		34	8.2	5.3	5.4	4.5

Discussion

Summary

How many and much?

- All Rs received at least one edit message
- Non-response edit messages were almost 70%

Which items?

- Items related to economic activity subject
- Items requiring computer skills like searching
- Numerical entries items: date, time, amounts

Summary

Respondents' Behaviors toward Edit

- 16.5% of edits was failed to be corrected on the first attempt
- Displayed Maximum 4th message on the specific items
- Detected satisficing behaviors among success

Classifying Satisficing Behaviors

- Reducing number of questions using branching questions
- Avoiding open-ended answer in other-specify questions
- Answering approximate value

Summary

Factors of Affecting Edit Message

- The number of edit messages is higher among aged and mixed response group

Limitation

Eye tracking test did not focus solely on edits, but whole process

- In some cases, edits were not fully functioned

The results could not be tested for statistical significance

- Used small number of voluntary samples
- Independent variables like response patterns were not assigned randomly

Implication

From the view point of questionnaire

- Avoid Complex navigations and functions (like searching)

From the view point of edit message

- Write clearly what the problem is and how to fix it

자료2 : Record Linkage Method for the Social Survey in Statistics Korea(정미옥 주무관, 7.15.)

Record Linkage Method for the Social Survey in Statistics Korea

Miock Jeong · Pilkeun Choi

Methodology Division, Statistical Research Institute



<http://sri.kostat.go.kr>

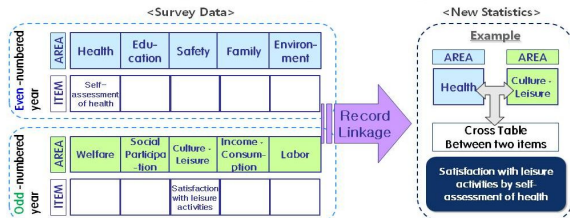
1. Introduction

Motivation

- Statistics Korea conducts the Social Survey
 - It provides information on subjective opinions related to people's quality of life
 - It includes 5 topics in every year and covers a total of 10 topics for 2 years
- It is not possible to analyze cross-tabulation across all topics of the Social Survey

Subject

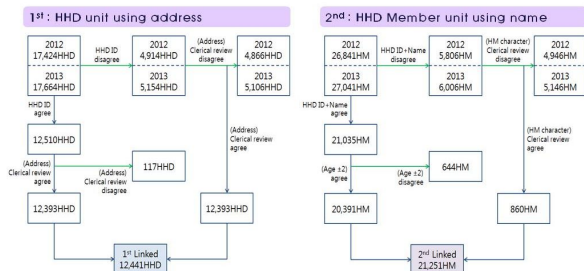
- Construct a new dataset through the linking of records from the year 2012-2013 of Social Surveys
- Two different linkage methods are considered :
 - the "deterministic" linkage method and the "probabilistic" linkage method



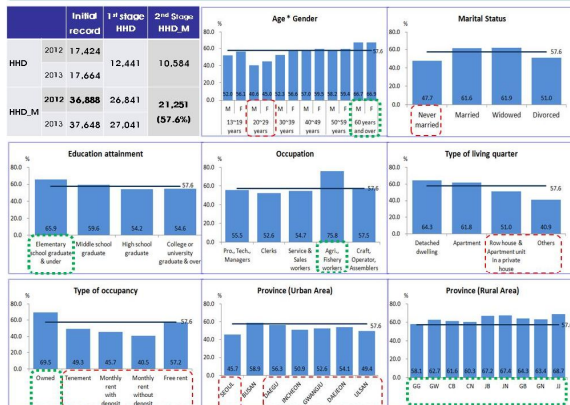
2. Deterministic Record Linkage

Linkage Process : 2 stage

- The data : 2012 - 2013 Social Survey file
 - Sampling Frame : 2010 Census Enumeration District (period of Sampling : 5 years)
 - Sampling Unit (Household), Survey Unit (Household member aged 13 and over)



Linkage Rate



- Linkage rates show some patterns: Records are not randomly linked.

Comparison of distribution between initial data and exact linked data

Character		Initial Data	Linked Data	Chi-Square	P-value
Gender	Male	46.6		7.13	0.007
	Female	52.4	53.5		
Age	13-19 years	10.8	10.1	198.35	<.0001
	20-29 years	11.2	8.3		
	30-39 years	16.9	16.0		
	40-49 years	19.9	20.1		
	50-59 years	17.3	17.7		
	60 years & over	23.9	27.7		
	Never married	25.9	21.5		
Marital Status	Married	61.9	66.2	162.58	<.0001
	Widowed	8.5	9.1		
	Divorced	3.7	3.3		
Education	Elementary	21.8	24.9	88.11	<.0001
	Middle grad.	14.4	14.9		
	High school grad.	34.6	32.5		
	College & Univ.	29.2	27.7		
Type of Living Quarter	Detached dwelling	35.5	37.5	72.64	<.0001
	Apartment	48.5	48.7		
	Row house	6.1	5.6		
	Apt. in a house	8.2	6.9		
	Others	1.7	1.3		
Type of Occupancy	Owned	65.6	72.2	294.52	<.0001
	Tenement	15.9	13.1		
	Monthly with dps.	9.5	7.6		
	Monthly no dps.	5.9	4.1		
	Free rent	3.2	3.0		

- The Linked data seems to be a different distribution from the initial data, so we should perform a probabilistic record linkage using unlinked records to preserve the initial distribution

3. Probabilistic Record Linkage

Linkage method : nearest neighbor random hot deck

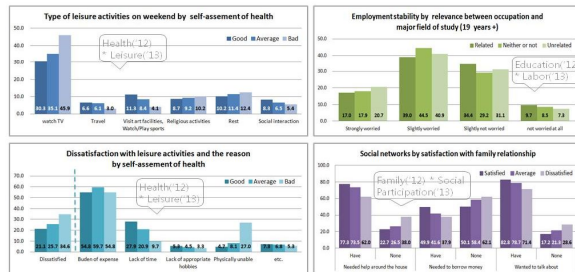
- Comparison of value of a particular variable from each record pair within a block
- For each record pair comparison, field weights are summed to form a linkage weight
- The one which gets the highest linkage weight is linked

- $m\text{-prob.} = P(\text{fields agree} | \text{records belong to same entity})$
- calculated by fields agreement ratio from deterministic linked data
- $u\text{-prob.} = P(\text{fields agree} | \text{records belong to different entity})$
- the number of categories from linking variables
- Agreement field weight = $\log_2(m/u)$: positive
- Disagreement field weight = $\log_2((1-m)/(1-u))$: negative

Item	Role	m-prob.	u-prob.	Agreement field weight	Disagreement field weight
Gender	2 Blocking var.				
AGE	14 Blocking var.				
Area	25 Blocking var.				
Marital status	4 Linking var.	0.972	0.250	1.96	-4.74
Education attainment	5 Linking var.	0.799	0.200	2.00	-1.99
Active status	2 Linking var.	0.828	0.500	0.73	-1.54
Occupation	6 Linking var.	0.766	0.167	2.20	-1.83
Status of worker	4 Linking var.	0.849	0.250	1.76	-2.31
Type of occupancy	5 Linking var.	0.856	0.200	2.10	-2.47

Results

- A new linked dataset is constructed by implementing these two methods in a sequence
 - All records(36,888) are included in a linked dataset
- Analysis of cross-tabulation across the items of two different years of the Social Survey.



4. Concluding Remarks

- This augmented dataset allows researchers to better understand humans and society, and also provides enhanced information for making social development policies.
- Furthermore, it can lead to savings in cost, time and respondent burden because it uses the information already collected without conducting additional surveys.
- A simulation study on the validity of the probabilistic record linkage results is in progress.



6th Conference of European Survey Research Association
Feel free to contact: mjeong@kostat.go.kr

【붙임2】 컨퍼런스 전경

□ 전시장



□ 일반세션 & 발표1



□ 포스터세션 & 발표2

