

Using the Theory of Socially Distributed Cognition to Analyse the Establishment Survey Response Process

Boris Lorenc

Statistics Sweden

The views expressed here are those of the author and do not necessarily reflect the views of Statistics Sweden

Content

Introduction

Socially Distributed Cognition (SDC) – ideas and principles

Propagation of representational states

Two studies that are motivated by, and grounded in, the principles of SDC

- Study I: Reporting about averages, an empirical study
- Study II: Assessing data availability

Consequences and conclusions



Introduction

Goal to understand the response process in establishment surveys

A cognitive approach, in the footsteps of CASM

But, on the way, a new unit of cognitive analysis is introduced, together with a new way to analyse the response process

Has consequences for practical work (measurement) and theoretical work (modelling) for surveying enterprises



Socially Distributed Cognition (SDC)

A position that boundaries of a cognitive system are determined by processes and interactions related to some high-level recurring task

Not necessarily so that a cognitive system is contained within the head of an individual

Flight deck of an aircraft, a telephone call centre or that part of the enterprise that responds to a survey can be viewed as cognitive systems



Study I: Reporting about averages

A survey by the Ministry of Enterprise among SME about punctuality of payment of their invoices by their organisational customers

Several questions of the form:

- What percent of your invoices <have been paid after due-in date> in the previous 12 months
- On average, how many <days after due-in date have the late payments arrived> in the previous 12 months

Questions easy to understand, data exist in records, accessible to respondents, retrievable



Study I: Reporting about averages (cont'd)

What needs to happen if a correct response is to be recorded (provided it is not known):

- Go through all the invoices in the reference period
- Register the required property (whether or not X, number of days of Y) for each of them
- Sum the registered properties and divide by the number of invoices issue in the reference period

Assumption: failure to map correctly will manifest itself in substantial discrepancies between the results acquired by the above questions and a 'golden standard'



Study I: Reporting about averages (cont'd)

A follow-up study of 300 of the respondents, with the goal of verification of the previous data collection technique

Asked to randomly select 10 invoices and record facts about them (date issued, date due-in, date paid, date tax paid, legal actions to acquire a late payment)

Same variables as in the original study estimated on the basis of recorded (invoice-level) data

Study I: Reporting about averages (cont'd)

	ITEM	ORIGINAL STUDY	FOLLOW-UP
1.	Average agreed time until payment is due (days)	27	27
2.	Percent of invoices with agreed time until payment is due longer than 30 days	14	18
3.	Percent of invoices with agreed time until payment is due equal to 60 days or longer	3	4
4.	Percent of invoices whose payment was late (arrived after the due-in date)	14	50
5.	Percent of invoices whose payment was more than 10 days late	4	13
6.	Percent of invoices whose payment was more than 30 days late	1	3
7.	Average length of delay of delayed payments (days)	10	9



Study I: Reporting about averages (cont'd)

Hypotheses viz. a substantial difference on item 4

- a 'psychological' hypothesis
- an 'information system' hypothesis

Lack of difference on item 7 compatible with the difference on item 4, provided an exponential model and left-truncation

- no evidence that the mapping breaks down
- possibly related to the importance of the question for E's

The matters bear as much on individual cognition (perception, memory, decision making,...) as on the socially distributed response process methodology

Propagation of representational states (PRS)

Questionnaire item 1

PRS item 1

Totalt antal elever i skolan

År	Antal elever	
	Pojkar	Flickor
1	<input type="text"/>	<input type="text"/>
2	<input type="text"/>	<input type="text"/>
3	<input type="text"/>	<input type="text"/>
4	<input type="text"/>	<input type="text"/>
5	<input type="text"/>	<input type="text"/>
6	<input type="text"/>	<input type="text"/>
7	<input type="text"/>	<input type="text"/>
8	<input type="text"/>	<input type="text"/>
9	<input type="text"/>	<input type="text"/>
Totalt	<input type="text"/>	<input type="text"/>

Spara



from record onto screen or printout

from screen or printout onto respondent

from respondent onto form (paper or web)

from form onto survey database

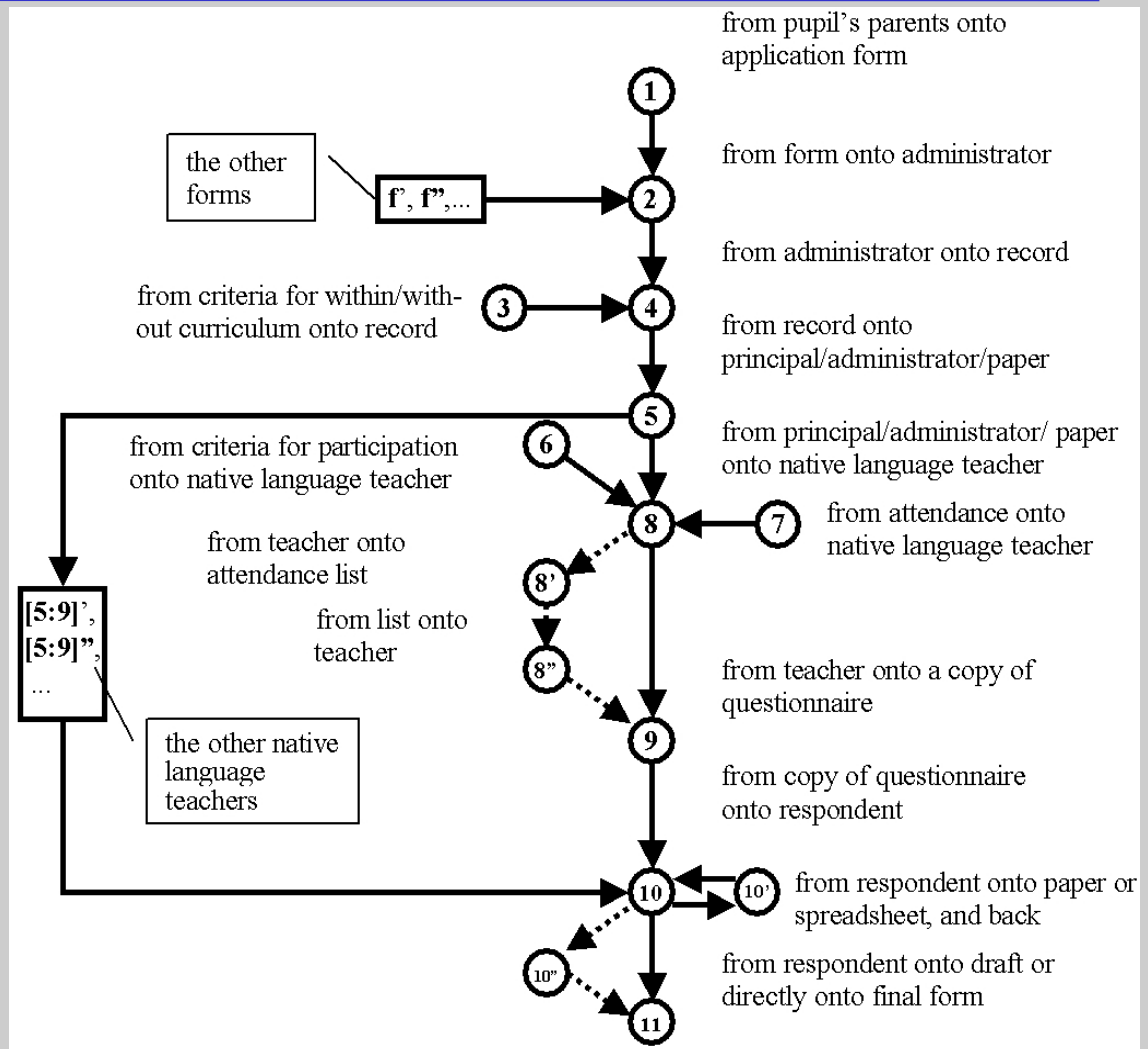
SCB PRS (cont'd)

Questionnaire item 2

År (1-9)	Modersmålet Lättaste sättet att söka fram till rätt språk, är att ange språkets första bokstav	Antal elever									
		berättigade till undervisning i modersmål som ämne		som deltar i undervisn. i modersmål som ämne och/eller studiehandledning på modersmålet		<i>därav</i> utanför timplane-bunden tid		som deltar i undervisning i svenska som andraspråk (SVA)			
		Pojkar	Flickor	Pojkar	Flickor	Pojkar	Flickor	Pojkar	Flickor		
1 ▾	Acholi ▾	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		
1 ▾	Acholi ▾	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		
1 ▾	Acholi ▾	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		
1 ▾	Acholi ▾	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		
1 ▾	Acholi ▾	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		

PRS (cont'd)

PRS item 2





PRB (cont'd)

PRS provides a way of formalising, analysing and displaying parts of the response process

In the analysis, e.g. probabilities of successful mapping may be attached to each arc, and also probabilities modelled for the whole graph




Study II: Assessing data availability

Results regarding surveys of individuals (Beatty & Herrmann, 2002) suggest that unavailability of data might lead to nonresponse (and possibly other quality problems like inaccuracy)

It is plausible to posit that something similar holds in establishment surveys

Thus, a methodological study initiated to identify levels of data availability, with the aim of using these levels in evaluation of ES questionnaire items



Study II: Assessing data availability (cont'd)

Here, 'unavailable' will mean: not able to correctly reach the producer's database

Continue to take a data perspective

A descriptive stance taken, as the goal is questionnaire item evaluation

Based on theoretical (Beatty & Herrmann, 2002; Willimack & Nichols, 2001; Willimack, Nichols & Sudman, 2002) and empirical evidence, the following framework is proposed:

Study II: Assessing data availability (cont'd)

For each item, determine:

	Media for propagation		Operations on representations, performed by the respondent
1.	Electronic	1.	None (solely digital processing)
2.	Respondent	2.	Re-write
3.	Electronic+respondent	3.	Arithmetic operations
4.	Respondent _b +respondent	4.	Estimation (on at least one component, if there is more than one)
5.	Electronic+respondent _b +respondent	5.	Guessing (on at least one component, if there is more than one)
6.	Other way of data collection+respondent	6.	Satisficing
7.	Other way of data collection+respondent _b +respondent		
8.	None		

Study II: Assessing data availability (cont'd)

Envisioned use: description of the response process on the item level

- Field use: based on direct empirical data (observations, interviews,..., collected by visiting enterprises)
 - may be aggregated later, for generalisation
 - necessarily expensive
- Use through proxies (experts): what is the likely response process given conditions commonly met in <this> kind of establishments
 - already aggregated
 - less expensive

Implication that processes further down on the two scales are more likely to lack quality

Conclusions

- SDC, like other theories, provides a way of looking at things
- It, like other theories, brings some aspects into focus and neglects others
- Best viewed as a complement to the other approaches (e.g. the hybrid model), giving new contributions to understanding of the response process
- Prominently, it brings into focus the path that information passes (or doesn't pass) in order to be correctly recorded in the producer's data base
- It also provides a theoretical place for practices established for survey participation (not mentioned here)
- It is a 'unified' theory and with representations that are more observable than what is the case in usual cognitive science



Acknowledgements

I thank Lilli Japiec, who suggested the line of research that led to the model for assessing data availability, as well as to colleagues that gave feedback and otherwise contributed to the studies presented here: Anette Björnram, Anna-Greta Erikson, Johan Erikson, Eva Furubjelke, Lisa Levinsson, Reijo Ljungberg and Andreas Persson.

Thank You

Boris Lorenc

Statistics Sweden
R&D Department
701 89 Örebro
Sweden

E-mail: boris . lorenc at scb . se