The impact of open-ended questions: A Multivariate Study of Respondent Engagement

Steven Gittelman, Ph.D.

Mktg, Inc.

New York

"It's the length, stupid!"

Kees de Jong

Research World, June 2010

Can anything be that simple?

If it is then creating a predictive model of engagement should be an easy task.

"There is no silver bullet.": Market Tools "Survey Score"

- A study of over a thousand studies.

- All from one panel source.

- "…survey design directly influences respondent…. engagement, in a consistent way."

- " …survey length proved to be generally predictive of most respondent engagement measures, there was wide variation in the design variables that were most influential in driving various measures of engagement."

Does it work?
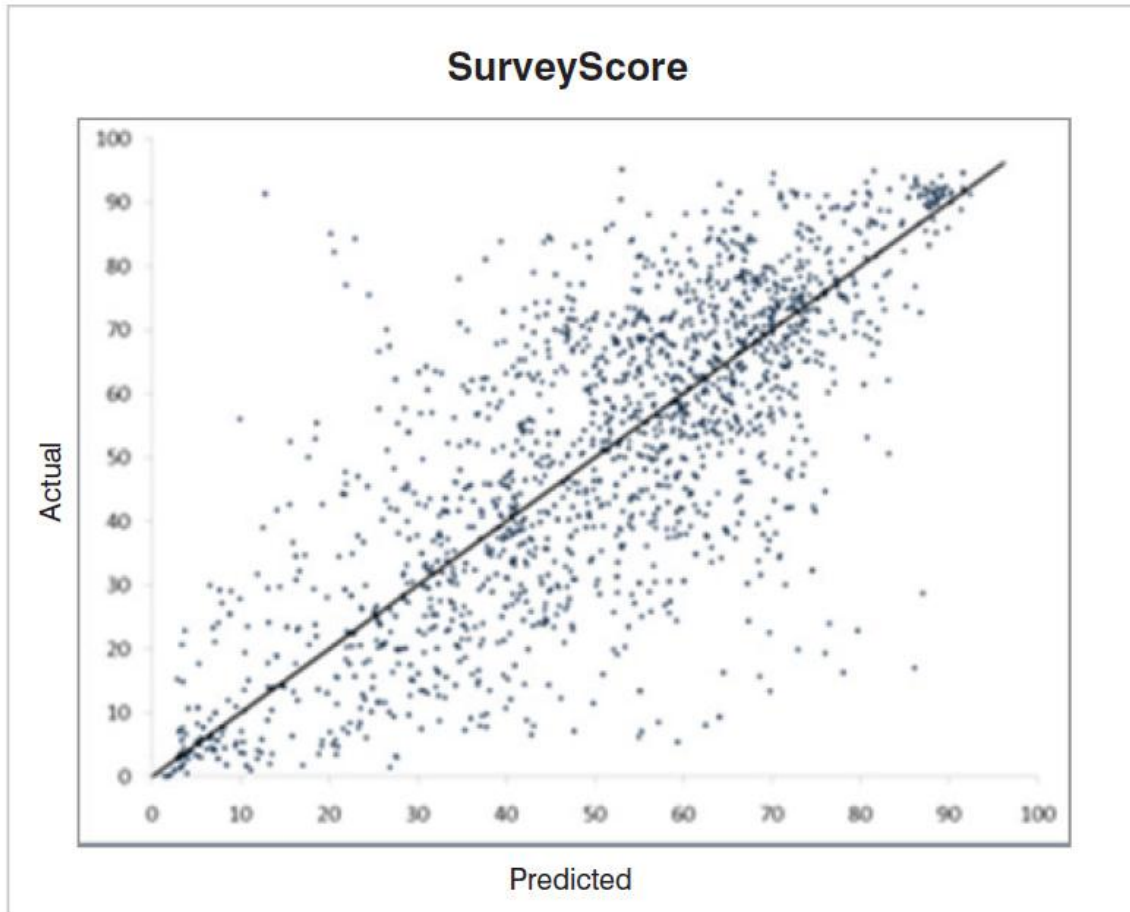
SurveyScore

Methods

- 1010 surveys consisting of >100 respondents.
- Excluding mall, B2B and physician studies.
- Multiple (20) panels, topics, and screening methods.
- A large number of questionnaire design variables.
- Four (total and partial straight lining, speeding, break offs) engagement variables.
- A large number of product/service categories.

Definitions

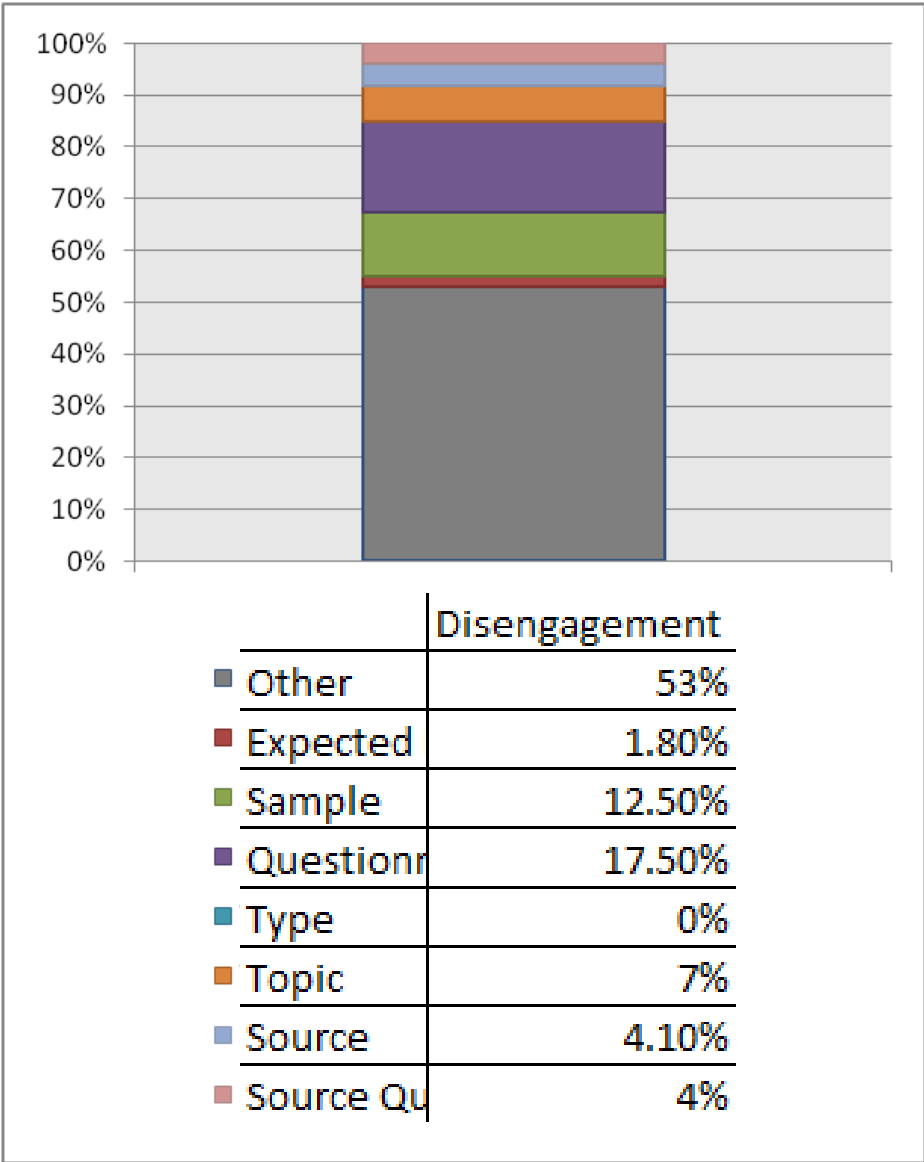**Disengagement** – break-offs, straight-liners, and speeders within a given survey.

Straight Lining – similar answers across multiple items within grid questions (<1 Standard Deviation of variance).

Sample and Demography –sample characteristics, screening methods, and topic.

Questionnaire structure-- length and % proportions of different types of questions (factual vs. opinion, single punch, etc.).

% of Disengagement Explained

	Disengagement
Other	53%
Expected	1.80%
Sample	12.50%
Questionn	17.50%
Type	0%
Topic	7%
Source	4.10%
Source Qu	4%

% of Straight-lining Explained

	Straightlining
Other	41.50%
Expected	2.90%
Sample	4.50%
Questionr	34.10%
Type	3.30%
Topic	3.10%
Source	5.70%
Source Qu	4.80%
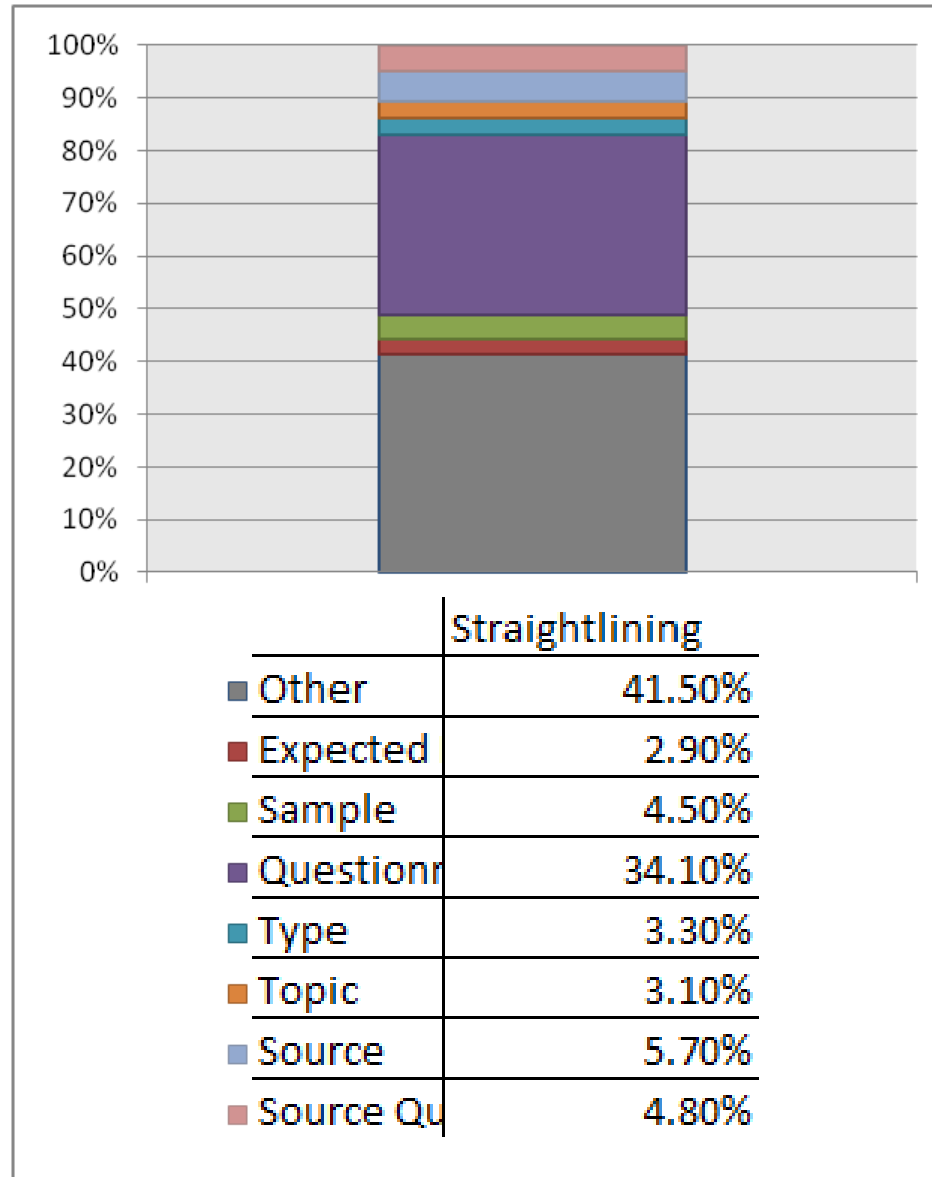
% of Disengagement Explained Within Sources
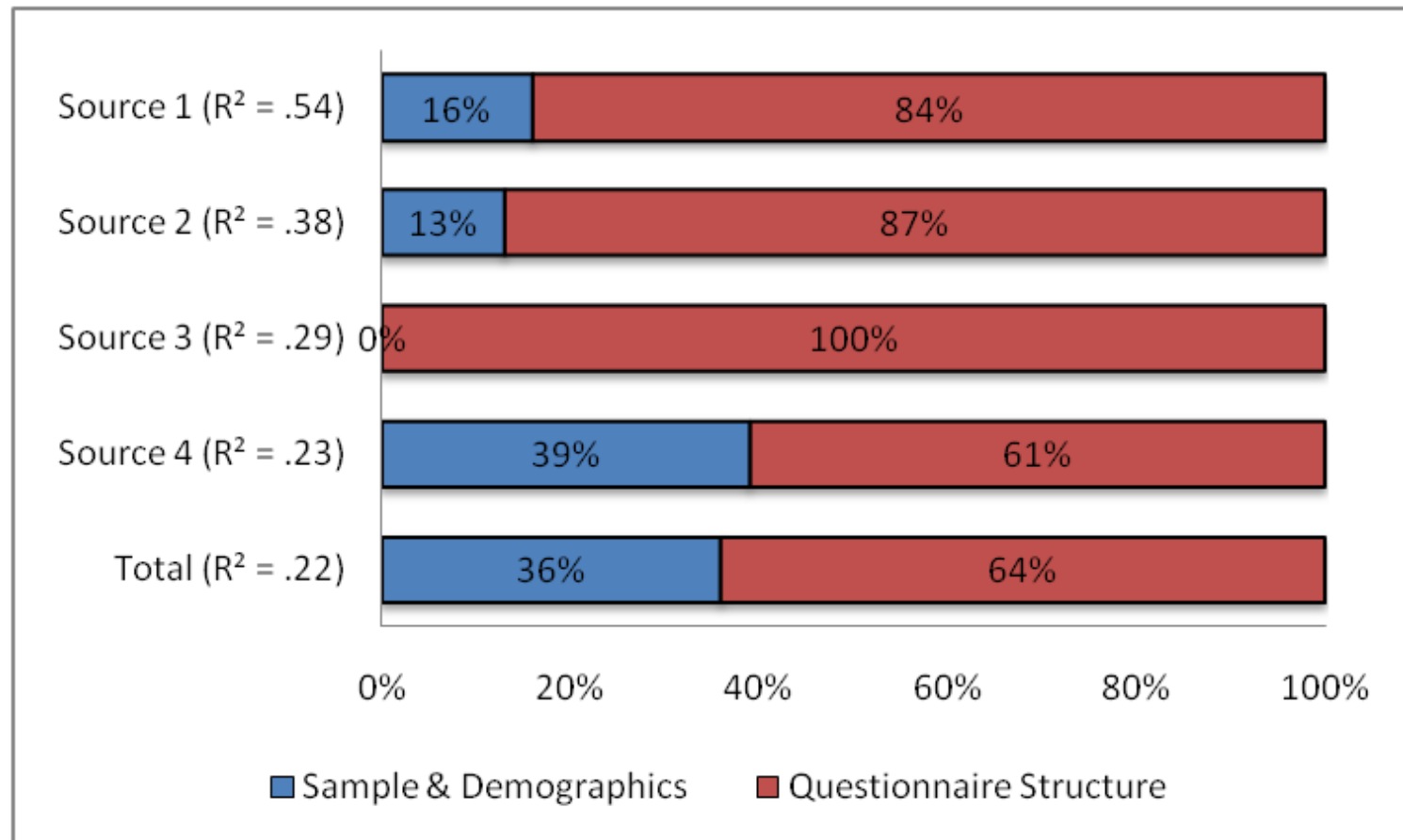
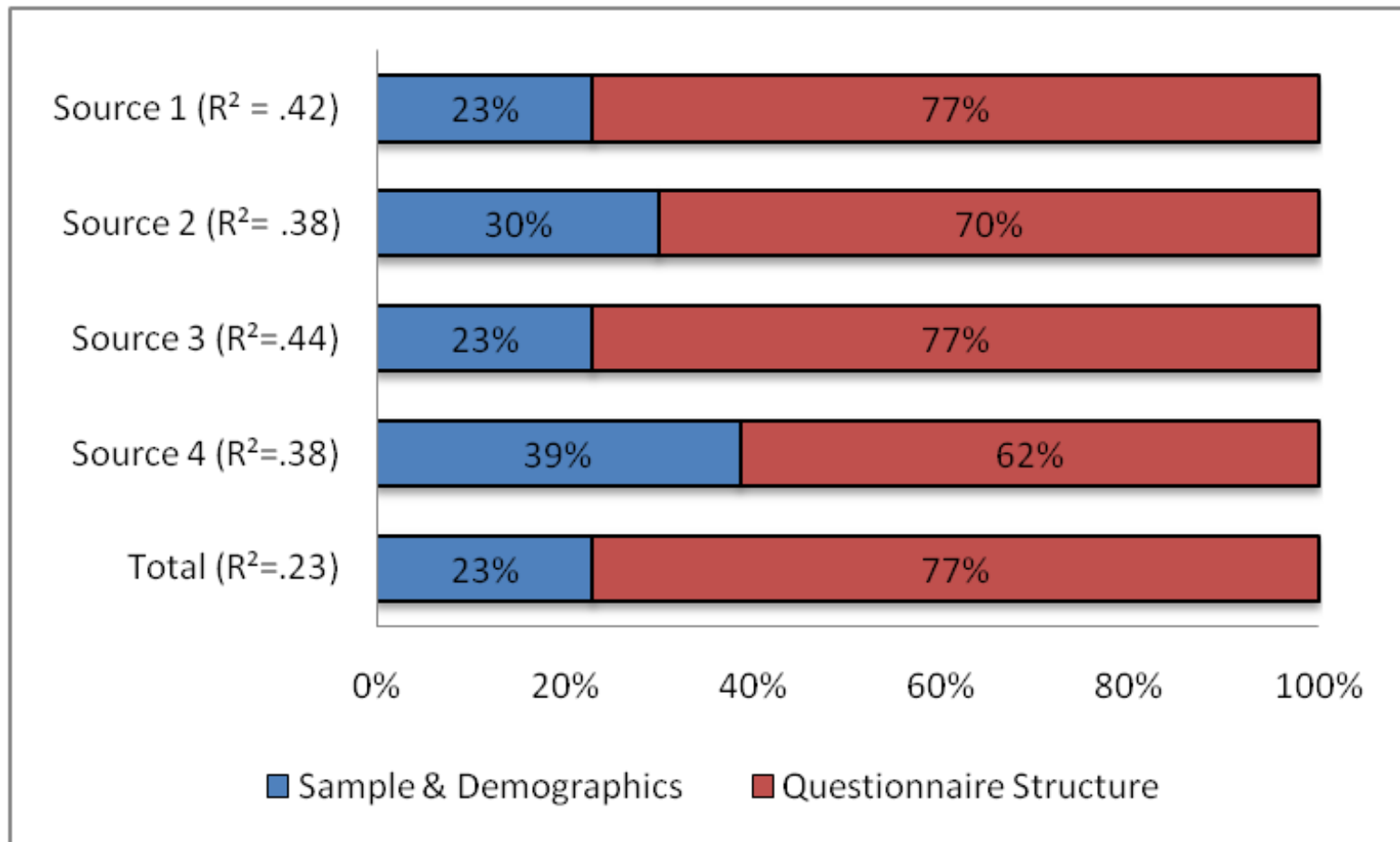# % of Straight-lining Explained Within Sources

Regression Model - Disengagement

		Source 1 (R^2 = .53)	Source 2 (R^2 = .38)	Source 3 (R^2 = .29)	Source 4 (R^2 = .23)	Total (R^2 = .12%)
	Number	78	87	224	158	1014
Sample & Demography	% Terminated	-	-	-	[red]	[red]
	Cropduster	-	-	-	-	-
	Concept Test	-	-	-	-	-
	Screened	-	-	-	[orange]	[orange]
	Gender	-	-	-	-	[blue]
	Age	[green]	-	-	-	-
	Income	[orange]	[blue]	-	-	-
Questionnaire Structure	Length	[red]	[red]	[red]	-	[orange]
	Question Length	[red]	[blue]	-	-	[orange]
	Early Demographics	-	-	-	-	-
	Graphics	[orange]	-	-	-	-
	Factual Ratio	[red]	[red]	[red]	-	[orange]
	Ranking Ratio	-	-	[blue]	-	[green]
	Adding Ratio	-	-	-	-	[orange]
	Reference Ratio	-	[red]	-	-	-
	Hypothetical Ratio	-	[blue]	[green]	-	-
	SP Ratio	-	-	-	[red]	[orange]
	MP Ratio	-	[red]	-	-	[orange]
	Numeric Ratio	[blue]	-	-	[green]	[orange]
	MNL Ratio	[blue]	-	-	-	-
	MNS Ratio	-	-	[red]	[red]	[orange]
	Open End Ratio	[red]	-	[red]	[red]	[red]

Legend:
- [red] >10%
- [orange] 0-10%
- [green] -10% - 0%
- [blue] < -10%

Regression Model - Straight-lining

		Source 1 (R² = .54)	Source 2 (R² = .38)	Source 3 (R² = .29)	Source 4 (R² = .23)	Total (R² = .12)
	Number	78	87	224	158	1014
Sample & Demography	% Terminated	-	-	[green]	[red]	[orange]
	Cropduster	-	[red]	-	[orange]	[orange]
	Concept Test	[red]	[red]	[orange]	-	[orange]
	Screened	-	-	-	-	-
	Gender	-	-	-	[orange]	[orange]
	Age	-	-	-	[orange]	-
	Income	[red]	-	[green]	-	-
Questionnaire Structure	Length	-	[red]	[red]	[green]	-
	Question Length	-	-	-	-	-
	Early Demographics	-	[green]	-	-	-
	Graphics	-	-	-	-	-
	Factual Ratio	[orange]	-	[red]	-	[orange]
	Ranking Ratio	-	-	-	-	[green]
	Adding Ratio	-	-	-	-	-
	Reference Ratio	-	[red]	-	-	-
	Hypothetical Ratio	-	-	-	[orange]	-
	SP Ratio	[red]	-	[red]	[red]	[red]
	MP Ratio	-	-	-	-	[orange]
	Numeric Ratio	[blue]	-	-	[blue]	-
	MNL Ratio	[blue]	[blue]	[blue]	-	[blue]
	MNS Ratio	-	-	[orange]	[blue]	[orange]
	Open End Ratio	[orange]	-	[red]	[orange]	[red]

Legend:
[red]	>10%
[orange]	0-10%
[green]	-10% - 0%
[blue]	< -10%

% of Incidence Explained – Clustered Regression
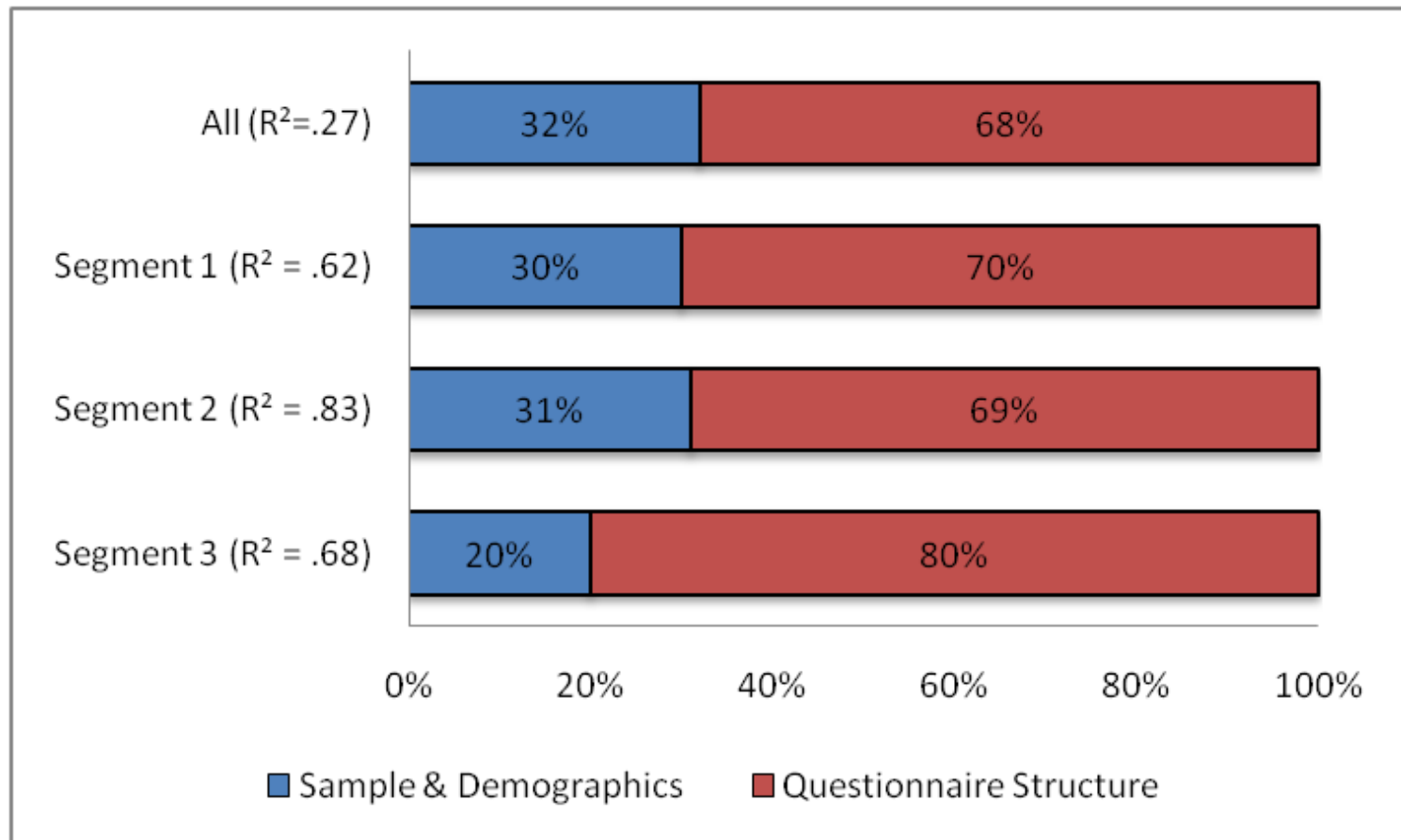
Clustered Regression Model – Breakoffs

		All (R² = .27)	Segment 1 (R² = .621)	Segment 2 (R² = .826)	Segment 3 (R² = .68)
	Number	1014	446	314	254
Sample & Demography	% Terminated	-	-	-	-
	Cropduster	green	green	green	orange
	Concept Test	-	-	orange	green
	Screened	orange	green	green	-
	Gender	blue	blue	green	orange
	Age	orange	orange	green	-
	Income	-	-	green	orange
Questionnaire Structure	Length	orange	orange	-	green
	Question Length	green	green	-	green
	Early Demographics	green	-	blue	-
	Graphics	-	green	-	-
	Factual Ratio	green	blue	-	green
	Ranking Ratio	green	-	green	-
	Adding Ratio	-	-	orange	blue
	Reference Ratio	-	-	orange	blue
	Hypothetical Ratio	-	-	green	-
	SP Ratio	orange	-	-	-
	MP Ratio	-	-	-	orange
	Numeric Ratio	orange	red	green	orange
	MNL Ratio	-	-	orange	blue
	MNS Ratio	orange	-	orange	-
	Open End Ratio	orange	-	green	-
Source Quality	Source Speeding	orange	red	red	-
	Source Straight-lining	blue	green	orange	blue
	Source Hyperactivity	orange	-	-	-
	Source Invalid	-	green	orange	green

Legend:

Color	Value
red	>10%
orange	0-10%
green	-10% - 0%
blue	< -10%

Clustered Regression Analysis - % Terminated
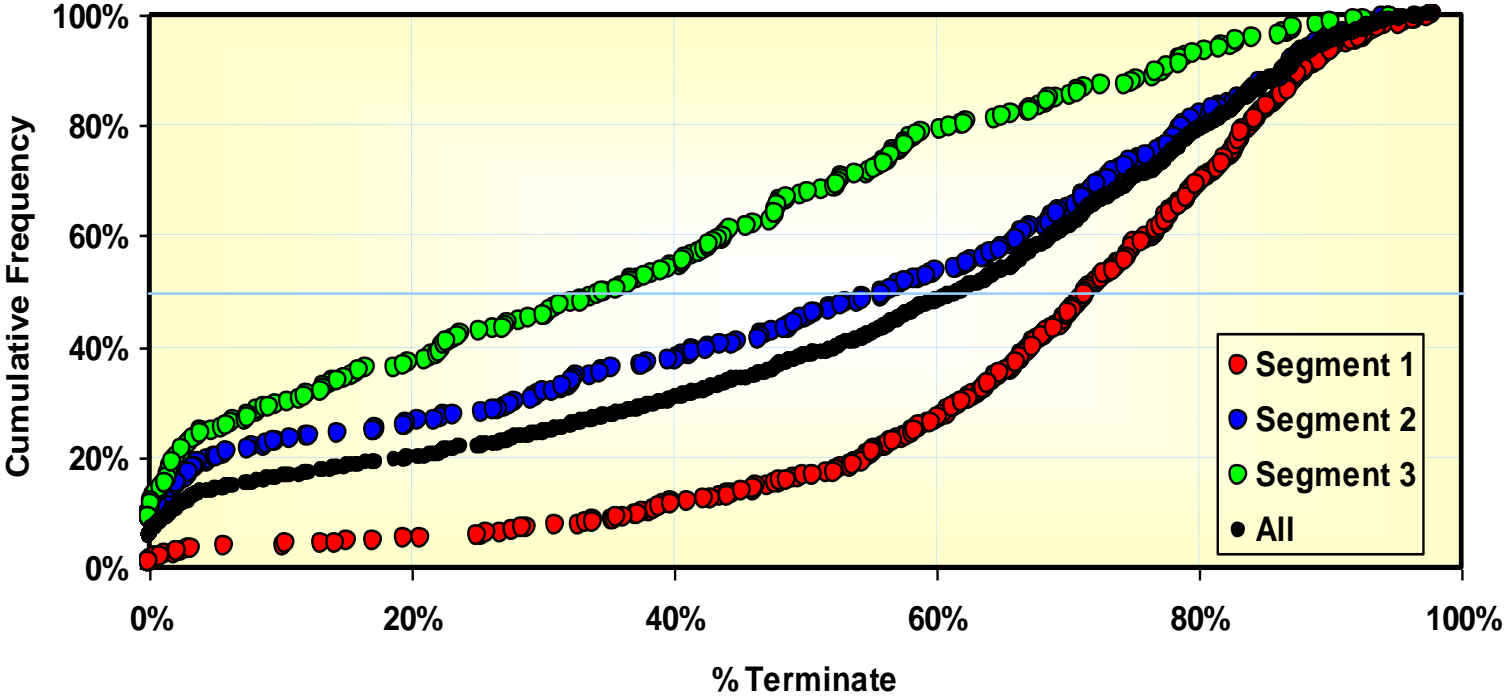
Regression Model – Disengagement by Incidence

		All (R² = .12)	Incidence >80% (R² = .21)	10%-80% Incidence (R² = .18)	Incidence < 10% (R² = .3)
	Number	1014	446	314	254
Sample & Demography	% Terminated	-	-	-	-
	Cropduster	🟧	-	-	-
	Concept Test	-	-	-	-
	Screened	🟧	🟥	🟩	🟦
	Gender	🟩	🟩	🟧	🟥
	Age	-	🟧	-	-
	Income	-	-	-	-
Questionnaire Structure	Length	🟧	🟩	-	-
	Question Length	🟥	🟧	🟧	-
	Early Demographics	-	-	-	🟥
	Graphics	-	-	-	🟦
	Factual Ratio	🟥	-	🟧	🟥
	Ranking Ratio	🟩	-	🟩	-
	Adding Ratio	🟧	🟥	-	🟧
	Reference Ratio	-	-	-	-
	Hypothetical Ratio	-	-	-	🟩
	SP Ratio	-	-	🟧	-
	MP Ratio	-	-	-	-
	Numeric Ratio	-	-	-	-
	MNL Ratio	🟩	-	🟩	-
	MNS Ratio	🟧	-	🟧	-
	Open End Ratio	🟧	-	🟥	-
Source Quality	Source Speeding	-	🟧	-	-
	Source Straight-lining	-	🟥	-	🟦
	Source Hyperactivity	🟧	-	🟧	-
	Source Invalid	🟥	🟧	🟥	-

Legend:
Color	Value
🟥 Red	>10%
🟧 Orange	0-10%
🟩 Green	-10% - 0%
🟦 Blue	< -10%

Conclusions

1. Driver's of engagement are complex.
2. Sourcing seems to have a strong influence: differences between sources are small but the driver's are different.
3. Length appears to be a good predictor but only within sources: it becomes inconsistent across sources.
4. Subject affinity appears to diminish disengagement.
5. Incidence may correlate with subject affinity by aggregating groups of similar demography or product interest.
6. Segments created through cluster regression show differences in their incidence profiles and the driver's of disengagement.
7. There is no silver bullet.

Questions?

Steven Gittelman

Steve@MktgInc.com | www.MktgInc.com
Blog: www.SampleMetrics.com