

Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey

Mirta Galesic¹

Dropouts can be a significant problem in web surveys, but theoretically motivated studies of this problem are rare. In this study, we use a dynamic theory of decision making, the decision field theory, to predict and explain behavior of respondents in a web survey. By registering respondents' momentary subjective experiences throughout the survey, we gained some insights into antecedents and consequences of respondents' decision to drop out. The results show that interest and experienced burden change throughout the survey, depending on characteristics of questions, respondents, and survey design. Respondents who drop out often express lower interest and higher experienced burden than the respondents who stay. Their growing preference for dropout can be detected in the decreased quality of their answers even before the point of dropout. The results could help in practical work and open new paths to theoretical explanations of survey behavior.

Key words: Web survey; dropouts; interest; burden; decision making; decision field theory.

1. Introduction

“Dropouts,” or the respondents who quit the survey prior to reaching its end, pose specific challenges to web survey research. In general invitation web surveys, which mostly recruit respondents through web-based advertisements, dropout rates – proportion of respondents who quit among all respondents – may be as high as 80 percent (O’Neil, Penrod, and Bornstein 2003; O’Neil and Penrod 2001), with an average of about 30 percent (Bosnjak and Tuten 2001; Lozar Manfreda and Vehovar 2002; Vehovar, Batagelj, Lozar Manfreda, and Zalatel 2002). For individually targeted web surveys, which use panels or other lists to directly recruit respondents, dropout rates are lower, but still average at about 15 percent (Lozar Manfreda and Vehovar 2002). In contrast, the proportion of premature terminations of face to face and telephone interviews is usually lower than five percent (e.g., Fricker, Galesic, Tourangeau, and Yan 2005; Groves and Kahn 1979; Stussman, Taylor, and Riddick 2003).

¹ Joint Program in Survey Methodology, University of Maryland, 1218 LeFrak Hall, College Park, MD 20742, U.S.A. Email: mgalesic@survey.umd.edu

Acknowledgments: This study was supported by the Croatian Ministry of Science, Project No. 0130406, PI: Branimir Sverko. The article was written during the author’s stay at the Joint Program in Survey Methodology, University of Maryland. Many thanks to Michael Bosnjak, Jerome Busemeyer, Mick Couper, Jill Dever, Frauke Kreuter, Darja Maslic-Sersic, Geraldine Mooney, Kenneth Rasinski, Hans-Joachim Schuetze, Branimir Sverko, Roger Tourangeau, the Associate Editor Annica Isaksson, and the anonymous reviewers for their helpful comments. The study was programmed using software package WARP-IT designed by RMplus, and conducted with the kind support of the research agency Puls, job-search website Posao.hr, and the Croatian Unemployment Bureau.

There are several studies that have explicitly dealt with the problem of dropouts in web surveys, yielding sometimes conflicting results. Some studies showed that prize draws may be effective in decreasing dropout rates (Bosnjak and Tuten 2003; Knapp and Heidingsfelder 1999; Musch and Reips 2000; O'Neil et al. 2003), though O'Neil and Penrod (2001) found opposing results. Asking personal information at the beginning of the survey or requesting manual input of login information might decrease dropouts later on (Frick, Baechtiger, and Reips 1999; Heerwegh and Loosveldt 2002a, 2003; but see O'Neil and Penrod 2001 and O'Neil et al. 2003). Open questions, matrix questions, longer loading times, and inappropriate visual design may contribute to higher dropout rates (Crawford, Couper, and Lamias 2001; Dillman, Tortora, Conrad, and Bowker 1998; Heerwegh and Loosveldt 2002b; Knapp and Heidingsfelder 1999). Respondents who are more interested in survey topics may be less prone to dropping out (such as the more frequent website visitors in the website evaluation survey described by Knapp and Heidingsfelder 1999).

The existing studies on dropouts provide useful information on correlates of dropouts, but they may be complemented in at least two ways. First, none of the studies manipulated order of the questions, which makes it hard to understand whether a dropout was caused by a particular question or by the cumulative effect of all previous questions in a questionnaire. Second, although some of the studies are designed to manipulate perceived burden of the respondents, they do not examine the actual respondents' subjective experience with the questionnaire (for an exception, see Crawford et al. 2001). Questions about interest and invested effort are sometimes asked at the end of web questionnaires (e.g., in Conrad, Couper, Tourangeau, and Peytchev 2005; and in Krosnick et al. 2002), but that is inappropriate for a study of dropouts. While in interviewer-administered surveys it is possible to observe respondents' reactions throughout the survey process, no one so far has investigated how respondents' subjective experience changes *during* a self-administered survey process and how it is related to dropout and other aspects of survey behavior.

A theory that may help in understanding how changes in respondents' subjective experience during a survey affect their behavior is *decision field theory*, a dynamic theory of decision making developed by Busemeyer and Townsend (1993) and Townsend and Busemeyer (1995). As opposed to classic expected utility theory, which is deterministic (it assumes that the preference relation is either "true" or "false") and static (disregards changes in preference during the deliberation), the decision field theory is probabilistic (allowing decisions to be made in the whole range from "true" to "false") and dynamic (allowing for changes in preferences as a function of deliberation time). The theory was successfully applied to a wide range of decision situations (e.g., Busemeyer and Diderich 2002; Finn 2002; Raab 2002; Stewart, Chater, Stott, and Reimers 2003). For explaining survey behavior, one of the most important concepts of this theory is the *inhibitory threshold*, the point which determines when the difference in the preference for one or the other action is large enough to provoke behavior. As long as the preference for one action is larger than that for the other, but not enough to cross the threshold, a person shows only an inclination towards the preferred activity but is not actually performing it.

In a survey situation, respondents often choose between two alternative actions – to continue or not to continue answering questions. At the beginning of a survey, while

factors that influenced their initial decision to participate (for example, incentives, short announced length, or general interest in the topic) are still influential, their preference to continue is higher than their preference to give up. But as the survey continues, the influence of negative aspects of participation (such as fatigue and boredom) becomes stronger and so does the preference to stop participating. However, until this change crosses their inhibitory threshold and affects behavior, respondents will continue to participate despite the growing tendency to stop. The preference to complete the questions may nevertheless decline, and that can result in lower quality of their answers.

Two aspects of respondents' subjective experiences may be especially important in studying survey behavior: interest in the questions, and burden experienced while answering them. *Interest* in a survey topic in general, and in the topic of each individual question in particular, can be a significant determinant of respondents' motivation to participate. In the absence of other positive aspects of participation, interest in the survey topic can significantly contribute to the respondent's initial decision to take part in a survey (Baumgartner and Rathbun 1997; Groves, Singer, and Corning 2000; Groves, Presser, and Dipko 2004) and possibly also their decision to stay (Knapp and Heidingsfelder 1999). *Burden* that respondents experience during participation is another factor that can influence their motivation. Everything else being equal, the more burdensome the questions are – whether because of their format or because of their content – the less motivated should respondents be to answer them. The effect of burden might be cumulative, whereby the burden experienced at each question is a function of both specific characteristics of that question and burden experienced while answering the preceding questions.

This study was designed to shed some light on the dynamics of respondents' behavior in a self-administered survey. First, we will observe incidence of dropouts in a web survey in which we experimentally manipulated announced length and type of incentives. We hypothesize that dropouts will be lower when the questionnaire is shorter and when some incentives are offered, since both of these factors should support respondents' preference for participation.

Second, we will establish whether dropouts in our web survey can be attributed to changes of interest and burden experienced while answering survey questions, beyond and above what can be explained by the announced length, incentives, and characteristics of the respondents. We predict that the overall interest in questions will be negatively, and overall experienced burden will be positively, related to the risk of dropout.

Third, we will examine how interest and burden change throughout the survey and how they differ for the respondents who drop out and those who stay until the end of the survey. Our hypothesis is that momentary interest will drop, and momentary burden increase before the point of dropout for the partial respondents, while they will remain relatively stable for complete respondents.

Fourth, we will examine what are some of the factors that may affect momentary interest and burden. One important group of factors might be formal characteristics of questions, such as their position in the questionnaire, whether they are open or closed, and how many of them are on one page. Momentary interest and burden could also be affected by more general characteristics of survey design, such as incentives and announced length, and by characteristics of the respondents.

Fifth, we will test some specific predictions from decision field theory. For example, it can be inferred from Busemeyer and Townsend's (1993) theory that respondents' increasing preference for dropout will be reflected in decreasing quality of their answers. In other words, questions immediately preceding the point of dropout would get lower quality answers than other questions. A related hypothesis is that a very burdensome or a very enjoyable question will momentarily affect respondents' preference to complete the subsequent questions, and this will in turn affect the quality of their answers.

Finally, we will examine whether dropouts have practical significance, by analyzing the quality of answers and the substantive results of dropouts compared to complete respondents. We will expand on our hypotheses more fully in the results section, after describing details of the study design and sample.

2. Study Design and Sample Characteristics

The study was conducted as a part of a larger, unrestricted, banner-advertised online survey on unemployment in Croatia. The screening part included ten questions on respondents' demographic characteristics and (un)employment status, and was the same for all respondents. The participants who stated that they were unemployed were allowed to proceed to the main part of the survey. This part consisted of approximately 180 questions divided in 20 blocks. Each question block occupied one web page and contained questions on a related topic – everyday activities, financial situation, importance of work, use of the Internet, physical and psychological health, social support, etc. Although the questionnaire was targeted towards the unemployed population, it included a wide range of topics, often not obviously related to unemployment.

To control for the effect of question position on tendency to drop out, the order of the question blocks (except for the screener questions and the first block) was *randomized*. At the end of each block the respondents were asked to evaluate how interesting they found the questions in that block (*interest*) and how much burden they experienced while answering them (*burden*), both on seven-point scales.² In this way, we collected measures of respondents' interest and experienced burden, independent of the question topic, throughout the survey. Unlike all of the other questions in this main part, the questions on interest and burden were mandatory. The respondents who opened all pages of the questionnaire ($n = 327$) were invited to comment on the survey. Approximately two thirds did so, and only four of them (1%) mentioned that they were annoyed by the repeated questions on interest and burden. It is of course possible that some of the respondents who dropped out before the last page did so because of the annoying interest and burden questions, and that this may be one of the reasons for the large overall dropout rate in our study (see below). None of the respondents mentioned anything related to the order of the questions.

The announced length of the survey and the motivation of the respondents were varied within a completely randomized experiment. The *three announced lengths* were 10, 20,

² The respondents were told that, since this was the first research on psychological consequences of unemployment in Croatia, their input on adequacy of the questions would be greatly appreciated. The exact text of the two questions was: "How interesting did you find the questions on this page?" and "How burdensome was answering the questions on this page?"

and 30 minutes. All three groups received the same questionnaire, but for ethical reasons the 10-minute group received – after the 5th block, or approximately one third of the survey – a message saying that the main part of the survey was now finished but that they could continue if they wished. The same message was shown to the 20-minute group after the 12th block, or approximately two thirds of the survey. The *different versions of motivation* offered to the respondents were “exclusive right to see the first results,” a chance to win 1,000 Kunas (approximately $\frac{1}{4}$ of the Croatian average monthly salary, or \$170), and a control version without any incentive.³ We included these variables in our analysis as controls.

A total of 2,339 participants started the survey; 958 stated that they were unemployed and were included in the main study. Demographic composition of the respondents is shown in Table 1. For the respondents who completed all questions, the average duration of the questionnaire that was announced as lasting 10 minutes was 14.4 minutes; the questionnaire that was announced as lasting 20 minutes took an average of 26.8 minutes to complete; and the questionnaire that was announced as lasting 30 minutes took an average of 38.4 minutes.

3. Results

3.1. Incidence of Dropouts

Overall, 41.8 percent of the respondents who started the survey dropped out before the end (Figure 1). Of the 380 respondents who started the 10 minute questionnaire, 31.8 percent dropped out before the 5th block (marking the end of the main part of the questionnaire in that group). Of the 285 respondents who started the 20 minute questionnaire, 43.2 percent gave up before the 12th block (which was the end of the main questionnaire in that group). In the 30 minute group, 53.2 percent of the 293 respondents who started that questionnaire dropped out before the end of the 20th block. In all three groups the dropout was largest at the beginning of the questionnaire: about 30 percent of all respondents gave up after approximately 10 minutes, or after approximately the first four pages of the questionnaire (Figure 1). After that, the dropout rate slowed down in all three groups. In fact, 27.6 percent of all respondents in the 10-minute group and 29.8 percent of those in the 20-minute group stayed for all 20 blocks, although they were informed that it was OK to stop after the 5th and 12th block, respectively.

3.2. Factors Affecting Dropouts

To test whether different characteristics of survey design, respondents’ demographics, and their experiences with the survey affect the risk of dropout, we used the Cox proportional hazards regression model. The dependent variable was the proportion of the total number of blocks in each time group. Of the survey design characteristics, the model included the two that affected all respondents: announced length and type of incentives (specific characteristics of individual blocks of questions were included in models of respondents’ survey experience, described in the next section). We hypothesized that longer announced

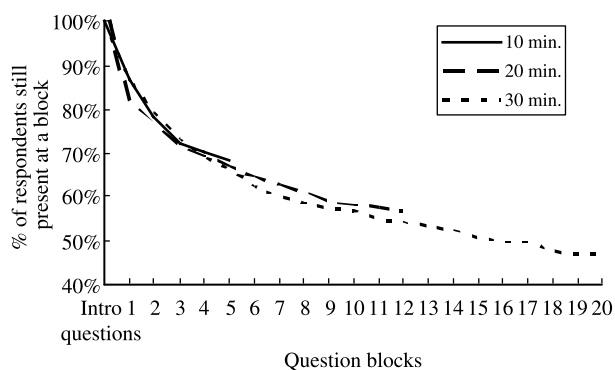
³ The study also included an immediate version of the prize draw, with the notification of the winner immediately after the questionnaire was completed. Those respondents were not included in this report; the results of that experiment are described in another paper (Tuten, Galesic, and Bosnjak 2004).

Table 1. Demographic composition of study respondents

	<i>n</i>	%
Gender		
Male	262	27.3
Female	696	72.7
Age		
Up to 24	281	29.3
25–34	474	49.5
35–44	138	14.4
45 +	65	6.8
Announced duration		
10 min.	380	39.7
20 min.	285	29.7
30 min.	293	30.6
Incentive		
No incentive	266	27.8
Exclusive results	339	35.4
Prize draw–delayed	353	36.8
Total	958	100.0

length (which corresponded to longer actual duration) will increase the risk of dropouts, and that some incentives (prize draw or study results) will produce fewer dropouts than no incentives at all.

Of the demographic characteristics, the model included respondents' gender, age, education, and level of work engagement. In line with the findings established in interviewer-administered surveys (e.g., Groves and Couper 1998), we hypothesized that women might be more cooperative, while older age and lower education might pose additional burden on participation and thus increase dropout. Although all respondents were officially unemployed, approximately one third of them engaged in some type of gray economy (occasionally or regularly working without a contract), or were occupied with household duties or studying. We hypothesized that the respondents with more work-related obligations would be less cooperative, i.e., would be more prone to dropout.



Note. After the screening questions, the 10 minute group got five blocks of questions in randomized order, the 20-minute group got 12 blocks, and the 30-minute group got 20 blocks of questions.

Fig. 1. Percentage of dropout at various parts of the survey, for 10, 20, and 30-minute groups

Finally, to account for subjective experience, the model included the average of respondents' estimates of momentary interest and experienced burden. We hypothesized that higher overall interest and lower overall experienced burden would contribute to lower risk of dropout. One potential obstacle in our analysis would be high correlation between evaluations of interest and burden, but the correlation was relatively moderate ($r = -0.20, p < .01$). Therefore we included both variables separately in our model.

The resulting Cox model is presented in Table 2. The model was statistically significant ($\chi^2(10) = 121.14, p < .001$), which means that the hypothesis that all coefficients in the model are zero could be rejected. The significance of each of the individual variables was tested using Wald statistic – the squared ratio of a coefficient (B) and its standard error (s.e.). In addition, we calculated the exponential value of a coefficient ($\text{Exp}(B)$), or hazard ratio, to describe the effect of a significant independent variable on the risk of dropout. Hazard ratios higher than 1 indicate that a variable increases the risk of dropout, while hazard ratios lower than 1 indicate that a variable decreases the risk of dropout when compared to the reference group (cf., Pratesi, Lozar Manfreda, Biffignandi, and Vehovar 2004).

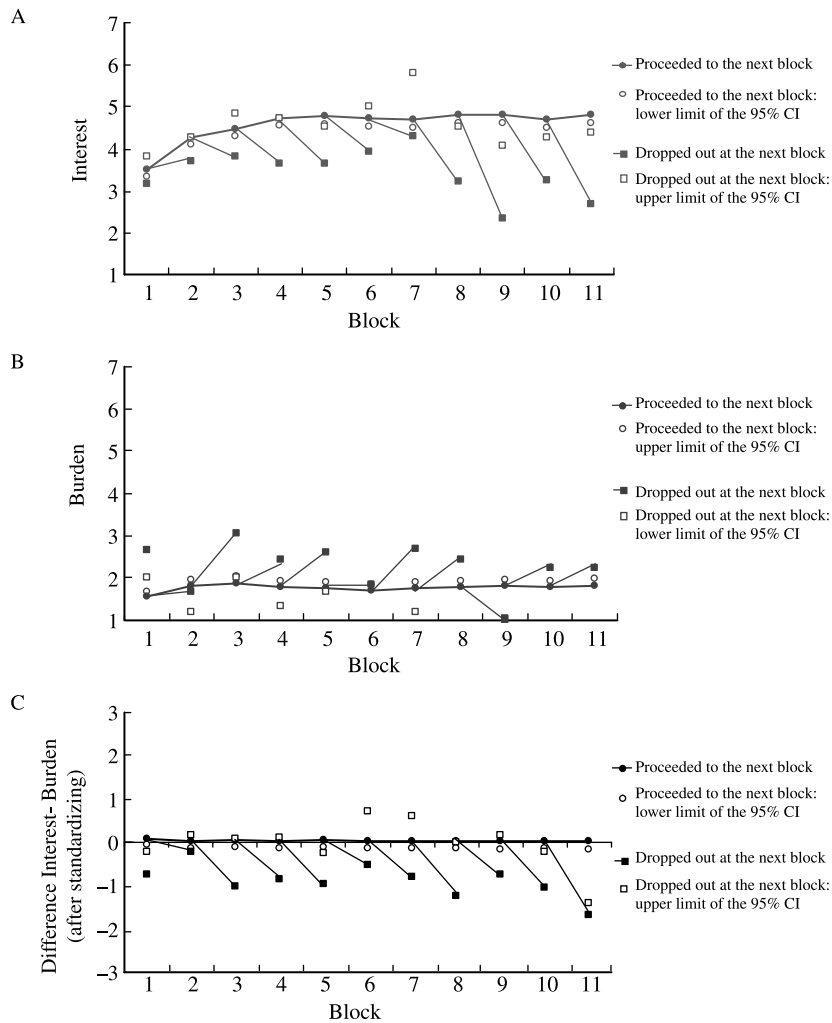
Announced length, respondents' age, as well as block-level interest and burden significantly affected the risk of dropout, while incentives, gender, education, and work-related education had no significant effects. Compared with the respondents who got the shortest questionnaire, the risk of dropout was approximately 20% higher for the participants who got the 20-minute questionnaire ($p = .01$) and approximately 40% higher for those who got the 30-minute questionnaire ($p < .01$). Contrary to our expectations, older respondents were somewhat more likely to complete the questionnaire – each year of age decreased the risk of dropout by approximately 1 percent ($p < .05$). But in line with our expectations, the subjective experience had a strong effect on the risk of dropout: the respondents with above-median interest had 40% lower risk of dropout than the respondents with below-median interest ($p < .01$). Similarly, the respondents with above-median experienced burden had 20% higher risk of dropout than the respondents with below-median experienced burden ($p < .01$).

Table 2. Effects of respondent- and survey design-related characteristics on risk of dropout: results of the Cox regression modeling

	B	S.E.	Wald	df	Sig.	$\text{Exp}(B)$
Announced time			17.61	2	0.00	
20 vs10 min	0.19	0.08	6.01	1	0.01	1.21
30 vs10 min	0.33	0.08	17.08	1	0.00	1.39
Incentives			0.48	2	0.79	
Results vs None	0.00	0.08	0.00	1	1.00	1.00
Prize draw vs None	0.05	0.08	0.33	1	0.57	1.05
Gender = Female	-0.08	0.07	1.18	1	0.28	0.92
Age in years	-0.01	0.00	4.78	1	0.03	0.99
Education = College	-0.08	0.07	1.42	1	0.23	0.93
Work obligations = Yes	-0.04	0.07	0.35	1	0.55	0.96
Above median interest	-0.56	0.07	71.62	1	0.00	0.57
Above median burden	0.21	0.07	10.09	1	0.00	1.24

3.3. Changes of Momentary Interest and Experienced Burden Throughout the Survey

We next look at the dynamics of change in respondents' subjective experiences during the survey. Figure 2 shows changes in momentary interest and burden experienced at each of the first eleven blocks, for the respondents from the 20- and 30-minute groups (those from the 10-minute groups were not asked to complete more than five blocks). For the respondents who did not drop out, interest in questions tended to increase at the beginning and then stabilize, while experienced burden increased slightly throughout the survey.



Sample sizes

<i>n</i> (proceeded)	449	416	399	383	366	355	345	336	331	325	321
<i>n</i> (dropped out)	35	30	17	14	17	12	10	9	3	8	3

Note: In Panels A and C, the difference between the respondents who proceeded and those who dropped out at a certain block is significant when \circ above \square ; in Panel B, the difference is significant when \square above \circ . In Panel C, interest and burden were standardized at the block level before calculating the difference.

Fig. 2. Changes in levels of interest and burden for the respondents who proceeded vs those who dropped out at each of the first 11 blocks (20- and 30-minute groups only)

For the respondents who dropped out at any later point in the survey, levels of interest were lower and levels of experienced burden higher already at the beginning of the questionnaire. Based on the estimates from all time groups, average initial interest was 3.31 for dropouts and 3.58 for those who stayed ($t(814) = 1.95, p = .05$), while the average initial burden was 1.77 for dropouts and 1.53 for those who stayed ($t(814) = -2.48, p = .01$). Furthermore, the interest of the respondents who gave up before the end tended to decrease, and experienced burden increase, just before the point of dropout.⁴ Consequently, the difference between the interest and burden experienced by the dropouts was always smaller than for the respondents who proceeded to the next block (Figure 2), significantly so in about 1/3 of the blocks. This is in line with the decision field theory, which would predict that the tendency for dropout would increase before the actual dropout happens.

3.4. Factors Affecting Interest and Experienced Burden

We now turn to analysis of the factors that may have affected the momentary interest and experienced burden. First, we compared formal characteristics of blocks whose difference between interest and burden was below median, with those that had above median interest-burden difference (the results were the same when we repeated the analysis with the ratio instead of with the difference between interest and burden). Blocks that were judged to be less interesting and more burdensome had more open questions (1.3 vs 0 open questions, $t(18) = 3.75, p < .01$) and longer average duration (140.9 vs 108.1 seconds, $t(18) = 2.98, p < .01$) than the blocks that were more interesting and less burdensome. Somewhat surprisingly, the two groups of blocks were not significantly different in total number of questions (6.9 vs 10.5 questions, $t(18) = 1.55, n.s.$).

To further examine the correlates of changes in momentary interest and burden we used hierarchical multilevel linear modeling (HMLM), a flexible tool for modeling multilevel data, especially for repeated measures models with missing data (Moskowitz and Hershberger 2002). The first level of our data included characteristics of question blocks: their formal characteristics (position, duration, number of open questions, and total number of questions in a block) and their subjective evaluation (block-related interest and experienced burden, both on a 7-point scale). The second level included data about respondents: their age in years, gender, and experimental group regarding announced duration and incentive.

The model of changes in interest throughout the survey can be very roughly captured by the following equations:⁵

$$\begin{aligned} \text{At Level 1 : Momentary interest} = & B_0 + B_1^*(\text{Block position}) + B_2^*(\text{Block duration}) \\ & + B_3^*(\text{Number of open questions in block}) \\ & + B_4^*(\text{Total number of questions in block}) + R \end{aligned}$$

⁴ The subjective estimates did not simply vary more among dropouts than among complete respondents. This was not so: average coefficient of variation for the evaluations of interest was 26.1 in both groups, and for the evaluation of burden it was 29.3 and 29.4 for dropouts and completes, respectively.

⁵ All predictors at Level 1 are continuous and grand-mean centered. All variables at Level 2 are dummy variables except for Age; Age is grand-mean centered and the other variables are left uncentered. Only the intercept computed at Level 1 was modeled in terms of Level 2 variables.

$$\begin{aligned} \text{At Level 2: } B_0 = & \gamma_{00} + \gamma_{01}^*(\text{Gender}) + \gamma_{02}^*(\text{Age}) + \gamma_{03}^*(\text{Announced duration} - 20\text{min.}) \\ & + \gamma_{04}^*(\text{Announced duration} - 30\text{min.}) + \gamma_{05}^*(\text{Incentive : Results}) \\ & + \gamma_{06}^*(\text{Incentive : Prize draw}); B_1 = \gamma_{10}; B_2 = \gamma_{20}; B_3 = \gamma_{30}; B_4 = \gamma_{40} \end{aligned}$$

In words, the momentary interest for a block i was modeled as a function of its position in the questionnaire, its duration in seconds, the number of open questions in the block, and the total number of questions in the block. For each of the respondents the characteristics of question blocks were regressed on their corresponding levels of interest (Level 1). The resulting regression slopes for each respondent were then, at Level 2, modeled in terms of respondents' gender and age and in terms of the announcements they received regarding the incentive for and the duration of the survey. A similar model was developed for changes in levels of momentarily experienced burden throughout the survey. The results are shown in Table 3.⁶

For both interest and burden, the intraclass correlation coefficient ρ suggested that using a multilevel model was appropriate, because question characteristics accounted for about 50% of the variability in interest and experienced burden among respondents. Relatively high values of coefficients λ for both models suggested that interest and burden were reliably measured on the respondents' level (Raudenbush and Bryk 2002).

Among characteristics of the question blocks (see Table 3), the significant predictors of momentary interest and burden were their position and duration, as well as the number of open questions they contained. Both interest and burden increased somewhat throughout the survey, which was reflected in the small positive coefficient of the block position. The number of open questions was negatively related to the level of interest – for each additional open question, the level of interest in this model dropped by 0.19 scale points. For each additional 100 seconds of block duration, the level of interest rose by 0.20 points, and the level of burden by 0.10 points. Total number of questions was negatively associated with burden, though on a very small scale (-0.07 points for 10 additional questions). One possible explanation of this somewhat surprising finding is that blocks with fewer questions more often than not included open questions.

Concerning the respondent-level characteristics, gender seemed to play an important role: women expressed much higher interest in the questions and reported significantly lower levels of burden. Although incentives seemed to increase the interest and lower the experienced burden, neither the promise of results nor the prize draw had a statistically significant effect. The longest announced duration (30 minutes) decreased the level of interest in questions by 0.30 points, but had no significant effect on experienced burden.

3.5. Further Predictions from the Decision Field Theory

The decision field theory would predict that the respondents who did not finish the survey could have been recognized even before they dropped out, because their preference for participation declined in one or more blocks preceding the point of dropout. To test this

⁶ We have also checked other models, but the results presented here appear to be very robust across somewhat different model specifications.

Table 3. Results of hierarchical multivariate linear modeling of changes in momentary estimates of interest and burden

	Interest		Burden	
	Coefficient	S.E.	Coefficient	S.E.
For intercept, β_{00} : Mean, γ_{00}	4.242**	0.153	2.035**	0.118
Gender = female, γ_{01}	0.368**	0.122	-0.210	0.099
Age in years, γ_{02}	0.006	0.006	-0.010	0.006
Announced time: 20 vs 10 min, γ_{03}	0.033	0.122	0.056	0.097
Announced time: 30 vs 10 min, γ_{04}	-0.299**	0.123	0.109	0.098
Incentive: Results vs None, γ_{05}	0.006	0.133	-0.095	0.105
Incentive: Prize draw vs None, γ_{06}	0.201	0.129	-0.154	0.101
Block position, β_{01}	0.031**	0.004	0.009**	0.003
Block duration in seconds, β_{02}	0.002**	0.0002	0.001**	0.0002
Number of open questions in the block, β_{03}	-0.193**	0.014	-0.018	0.01
Total number of questions in the block, β_{04}	-0.001	0.003	-0.007**	0.002
	$\rho = 0.538, \lambda = 0.876$		$\rho = 0.568, \lambda = 0.887$	

* $p < .05$; ** $p < .01$

prediction, we calculated the average item response rate for blocks that immediately preceded dropout vs blocks that did not, for the respondents who completed at least one of the random blocks but later dropped out ($n = 188$). For blocks not immediately preceding dropout, the average item response rate was 94 percent, while for blocks immediately preceding dropout the average item response rate was 89 percent. Even after accounting for clustering by respondents using SUDAAN, the difference was statistically significant ($t(187) = 2.22, p = .03$). Similar results were obtained for the length of answers to open questions (for 95 respondents who were assigned open questions and who dropped out at some point). For blocks not immediately preceding dropout, average length of answers to open questions was 94.8 characters, while for the blocks just before the dropout it was only 48.9 characters ($t(94) = 3.69, p < .01$).

Another prediction of the decision field theory is that an interesting or otherwise relatively enjoyable block should increase the preference for completing the following block; whilst a relatively burdensome block should decrease the preference for completing the following block. To test this hypothesis, we fitted two hierarchical linear models with block-level item response rate and length of answers to open questions as the dependent variables, and the difference between interest and burden experienced in the *previous* block, and block position, as the independent variables (details available from the author on request). Prior to calculating their difference, both interest and burden were standardized at the block level. As predicted, the larger the difference between interest and burden at the preceding block, the higher the item response rates in the subsequent block ($B = 0.023$,

$t(9142) = 6.01, p < .01$), and the longer the answers to open questions in the subsequent block ($B = 15.663, t(1885) = 2.46, p < .01$).

3.6. Differences Between Dropouts and Complete Respondents in Quality of Answers

Overall quality of data was lower for the respondents who dropped out (at any point of the survey) than for those who stayed until the end. The item response rate of an average block completed by the respondents who later dropped out was 91 percent, while the corresponding rate for those who stayed was 96 percent ($t(956) = 8.04, p < .01$). Similarly, the average length of answers to open questions was 46.9 characters for the dropouts and 105.4 for those who stayed ($t(677) = 8.00, p < .01$).

3.7. Differences Between Dropouts and Complete Respondents in Substantive Results

Were there any differences in substantive results between the respondents who dropped out and those who stayed until the end? To answer this question, we compared these two groups of respondents regarding their answers to 173 substantive closed questions, for each of the three questionnaire lengths. In the 10- and 20-minute questionnaires, the answers for the two groups differed at the 0.05 level for only 4.8 and 5.4 percent of the questions, respectively. This could have been expected by chance alone. In the 30-minute questionnaire, the two groups differed for 8.4 percent of the questions, slightly above chance. The differences disappeared, nevertheless, once we controlled for demographic variables (gender, age, education, workload) and incentives (prize draw, results, or none). Similarly, when we compared only the two extreme groups – the respondents who dropped out after four or fewer blocks (thus not completing even the shortest questionnaire) and the respondents who stayed for all 20 blocks of questions (independently of the announced length), we found only 8.1 percent significantly different substantive answers. These differences also mostly disappeared once we controlled for demographics and incentives.

4. Discussion and Conclusions

The methodology used in this study was innovative in a couple of ways. First, in order to avoid confounding of question position with its other characteristics that may have affected respondents' motivation, we randomized the order of blocks of thematically related questions. The resulting questionnaire was probably not as cohesive as a typical survey questionnaire, and this might have been one of the reasons for the high dropout rate we observed. Nevertheless, there was no reason to believe that any order of the question blocks appeared more or less disorganized than others. Second, respondents were reporting their momentary subjective experience (interest and experienced burden) throughout the questionnaire. While this procedure might have increased the overall dropout rate and is probably not recommendable for most practical studies, it enabled us to track the dynamics of respondents' motivation and its relations to indicators of data quality throughout the survey.

The results offer some insights into antecedents and consequences of a decision to drop out. Besides the announced length of the questionnaire, respondents' subjective experience during the survey significantly affected chances for dropout. The higher their overall interest in questions and the lower their overall experienced burden, the lower the

risk of their dropping out. In addition, respondents who dropped out after a certain block often reported significantly worse subjective experience with that block than the respondents who stayed.

Evaluations of interest and burden were related to certain characteristics of questions. For example, the more question blocks respondents had already completed, and the longer it took them to answer questions in a block, the higher was their experienced burden at that block. The more open questions there were in the block, the less interesting was the block perceived to be. Somewhat surprisingly, interest in a given block was also positively related to the time needed to complete it, as well as to its position in the questionnaire. This might suggest that interest in questions might balance the increasing burden experienced during a lengthy survey. On the other hand, subjective experience was not greatly affected by incentives; and only the longest announced length had a significant negative affect on interest in questions.

We also found support for two interesting predictions from the decision field theory. In line with this theory, it was possible to objectively observe respondents' tendency for dropout even before they actually gave up. In the block preceding the point of dropout, item response rates were significantly lower, and answers to open questions shorter than in blocks not followed by dropout. Another prediction was that respondents' preference to complete a block will depend on their subjective experiences with the preceding block. This was indeed found: when a block was preceded by a relatively enjoyable block, its item response rates were higher, and answers to open questions longer. In future studies, it might be worth investigating whether strategically placed enjoyable or at least easier questions can improve the overall quality of data gathered by otherwise burdensome questionnaires.

How much of a problem are dropouts? This study confirms previous findings that dropout rates in web surveys can be quite high. Our overall dropout rate was 42 percent, somewhat higher than the 30 percent rate that is usually observed in this type of web surveys. Nevertheless, the respondents who dropped out did not provide very different substantive answers than the respondents who stayed until the end. The fact that the order of question blocks was randomized certainly helped: each block appeared at all places in the survey, and dropouts were, on average, asked the same questions as the complete respondents. The differences in substantive results would obviously be larger if the respondents who dropped out were never asked certain questions.

On the other hand, we found systematic differences in the quality of answers of the respondents who gave up at some point as compared to those who stayed. The dropouts had a somewhat lower overall item response rate, and they gave significantly shorter answers to open questions. The quality of their data further deteriorated as they approached the point of dropout. This suggests that dropouts can be problematic in web surveys and care should be taken to prevent them.

What can be done to prevent dropouts? It is obviously important to keep the respondents' interest throughout the survey on a high level. Interest can motivate respondents to continue answering, and give answers of better quality, even as the burden increases with the duration of the questionnaire. The best way to keep respondents' subjective experience positive may not be to give incentives – at least in this study, they did not have much effect. Rather, one should try to make the actual questions as easy and enjoyable as possible. Open questions in particular seem to negatively affect interest in the questions. If the questionnaire must

include burdensome questions, it might be worth interspersing them at least occasionally with more enjoyable and easier questions, in order to boost the respondents' subjective experience and keep them in the survey.

This study has several limitations, and it should be considered as a preliminary examination of the feasibility of the dynamic approach to survey behavior. First, the mode of data collection was a self-administered web survey. The mechanisms that affect respondents' subjective experiences and cause dropout in interviewer-administered modes, such as telephone and personal ones, must be quite different. Characteristics of interviewers and their interaction with respondents certainly strongly affect respondents' willingness to complete an interview. Second, the survey was conducted on a specific population – of unemployed persons, who were probably more interested in this study than an average respondent is in a typical market or governmental survey. Third, the study was conducted in Croatia: it is possible, though no such data exist, that there are some cultural differences in the ways people respond to web surveys there as compared to other European countries and the U.S. Fourth, we attempted to measure only respondents' interest and burden, but there are many other dimensions of survey experience that can significantly influence their behaviors. The measures were very simple (single scales), prone to various sources of measurement error. On the other hand, even the use of such simple measures would be impossible in many situations due to limited time and resources. Although limited, the measures of interest and experienced burden correctly identified various patterns in respondents' behaviors.

In conclusion, respondents' behaviors in self-administered surveys may be significantly affected by changes in respondents' subjective experiences throughout the survey. In this study we tried to register the interest and burden the respondents experienced at each page of a web questionnaire. Independently of their content, pages that appeared later in the questionnaire, that contained more questions or simply open questions, and that required longer time to complete, negatively affected respondents' subjective experiences and were followed by dropout more often than other pages. In accord with the decision field theory (Busemeyer and Townsend 1993), respondents who dropped out showed decreased quality of answers to questions preceding the point of dropout, reflecting their rising tendency to drop out. Especially enjoyable items, on the other hand, tended to improve the quality of answers given to subsequent questions. Overall, the dynamic approach to survey behavior proved to be promising in predicting and explaining dropouts in web surveys.

5. References

- Baumgartner, R. and Rathbun, P. (1997). Prepaid Monetary Incentives and Mail Survey Response Rates. Paper presented at the annual conference of the American Association for Public Opinion Research, Norfolk, VA.
- Bosnjak, M. and Tuten, T.L. (2001). Classifying Response Behaviors in Web-based Surveys. *Journal of Computer-Mediated Communication* 6(3).
- Bosnjak, M. and Tuten, T.L. (2003). Prepaid and Promised Incentives in Web Surveys – An Experiment. *Social Science Computer Review*, 21, 208–217.
- Busemeyer, J.R. and Diderich, A. (2002). Survey of Decision Field Theory. *Mathematical Social Sciences*, 43, 345–370.

- Busemeyer, J.R. and Townsend, J.T. (1993). Decision Field Theory: A Dynamic-cognitive Approach to Decision-making in an Uncertain Environment. *Psychological Review*, 100, 432–459.
- Conrad, F.G., Couper, M.P., Tourangeau, R., and Peytchev, A. (2005). Effectiveness of Progress Indicators in Web Surveys: First Impressions Matter. *Proceedings of SIGCHI 2005: Human Factors in Computing Systems*, Portland, OR.
- Crawford, S.D., Couper, M.P., and Lamias, M.J. (2001). Web Surveys: Perception of Burden. *Social Science Computer Review*, 19, 146–162.
- Dillman, D.A., Tortora, R.D., Conradt, J., and Bowker, D. (1998). Influence of Plain versus Fancy Design on Response Rates for Web Surveys. Paper presented at the annual meeting of the American Statistical Association, Dallas, TX.
- Finn, P.R. (2002). Motivation, Working Memory, and Decision Making: A Cognitive-Motivational Theory of Personality Vulnerability to Alcoholism. *Behavioral and Cognitive Neuroscience Reviews*, 13, 183–205.
- Frick, A., Bächtiger, M.T., and Reips, U.-D. (1999). Financial Incentives, Personal Information and Drop-out in Online Studies. In U.-D. Reips and M. Bosnjak (eds), *Dimensions of Internet Science*, 209–220. Lengerich: Pabst Science Publishers.
- Fricke, S., Galesic, M., Tourangeau, R., and Yan, T. (2005). An Experimental Comparison of Web and Telephone Surveys. *Public Opinion Quarterly*, 69, 370–392.
- Groves, R.M. and Couper, M.P. (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley.
- Groves, R.M. and Kahn, R.L. (1979). *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- Groves, R.M., Presser, S., and Dipko, S. (2004). The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly*, 68, 2–31.
- Groves, R.M., Singer, E., and Corning, A. (2000). Leverage-saliency Theory of Survey Participation: Description and Illustration. *Public Opinion Quarterly*, 64, 299–308.
- Heerwegh, D. and Loosveldt, G. (2002a). Web Surveys. The Effect of Controlling Survey Access Using PIN Numbers. *Social Science Computer Review*, 20, 10–21.
- Heerwegh, D. and Loosveldt, G. (2002b). An Evaluation of the Effect of Response Formats on Data Quality in Web Surveys. *Social Science Computer Review*, 20, 471–484.
- Heerwegh, D. and Loosveldt, G. (2003). An Evaluation of the Semi-automatic Login Procedure. *Social Science Computer Review*, 21, 223–234.
- Knapp, F. and Heidingsfelder, M. (1999). Drop-out Analysis: Effects of the Survey Design. In U.-D. Reips and M. Bosnjak (eds), *Dimensions of Internet Science*, 221–230. Lengerich: Pabst Science Publishers.
- Krosnick, J.A., Holbrook, A.L., Berent, M.K., Carson, R.T., Hanemann, W.M., Kopp, R.J., Mitchell, R.C., Presser, S., Ruud, P.A., Smith, V.K., Moody, W.R., Green, M.C., and Conaway, M. (2002). The Impact of “No Opinion” Response Options on Data Quality: Non-attitude Reduction or an Invitation to Satisfice? *Public Opinion Quarterly*, 66, 371–403.
- Lozar Manfreda, K. and Vehovar, V. (2002). Survey Design Features Influencing Response Rates in Web Surveys. Paper presented at The International Conference on Improving Surveys, Copenhagen, Denmark.

- Moskowitz, D.S. and Hershberger, S.L. (eds) (2002). *Modelling Intraindividual Variability with Repeated Measures Data: Methods and Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Musch, J. and Reips, U.-D. (2000). A Brief History of Web Experimenting. In M.H. Birnbaum (ed.), *Psychological Experiments on the Internet*, 61–87. San Diego: Academic Press.
- O’Neil, K.M. and Penrod, S.D. (2001). Methodological Variables in Web-based Research That May Affect Results: Sample Type, Monetary Incentives, and Personal Information. *Behavior Research Methods, Instruments and Computers*, 33, 226–233.
- O’Neil, K.M., Penrod, S.D., and Bornstein, B.H. (2003). Web-based Research: Methodological Variables’ Effects on Dropout and Sample Characteristics. *Behavior Research Methods, Instruments, and Computers*, 35, 217–236.
- Pratesi, M., Lozar Manfreda, K., Biffignandi, S., and Vehovar, V. (2004). List-based Web Surveys: Quality, Timeliness, and Nonresponse in the Steps of the Participation Flow. *Journal of Official Statistics*, 20, 451–465.
- Raab, M. (2002). T-ECHO: Model of Decision Making to Explain Behaviour in Experiments and Simulations Under Time Pressure. *Psychology of Sport and Exercise*, 3, 151–171.
- Raudenbush, S.W. and Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Sage.
- Stewart, N., Chater, N., Stott, H.P., and Reimers, S. (2003). Prospect Relativity: How Choice Options Influence Decision Under Risk. *Journal of Experimental Psychology: General*, 132, 23–46.
- Stussman, B.J., Taylor, B.L., and Riddick, H. (2003). Partials and Break-offs in the National Health Interview Survey, 2002. Paper presented at the 2003 Federal Committee on Statistical Methodology Research Conference, Arlington, Virginia.
- Townsend, J.T. and Busemeyer, J. (1995). Dynamic Representation on Decision-making. In R.F. Port and T. van Gelder, *Mind as Motion: Explorations in the Dynamics of Cognition*, 101–120. Cambridge, MA: MIT Press.
- Tuten, T.L., Galesic, M., and Bosnjak, M. (2004). Effects of Immediate versus Delayed Notification of Prize Draw Results on Response Behavior in Web Surveys – An Experiment. *Social Science Computer Review*, 22, 377–384.
- Vehovar, V., Batagelj, Z., Lozar Manfreda, K., and Zalatel, M. (2002). Nonresponse in Web Surveys. In Groves, R., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (eds), *Survey Nonresponse*, 229–242. New York: John Wiley and Sons, Inc.

Received March 2004

Revised April 2006