stance, refer to the aggressive and competitive behaviour of political actors, conflicts, and power struggles, or the lack of considering different interests. The alternative treatment (Tx_2) consists of newspaper articles with negative information about the efficiency of political decision-making. Articles, for instance, refer to delays, protracted decisions, time-consuming procedures, a lack of efficiency of decision-making processes, and the indecisiveness of political actors (see Section 6.2.2). The control group does not get any treatment. The experimental design is a between subjects-design (cf. Grabe & Westley, 2003, p. 285) .

The target population of this study consists of citizens from the German-speaking part of Switzerland who are at least 18 years old and hold voting rights. The study's participants were recruited in collaboration with the Swiss online election information tool smartvote (www.smartvote.ch). The registered users of this platform regularly receive newsletters. One of those newsletters for users of the German-speaking population contained information on the planned study as well as contact information for readers interested in participating. The sample, hence, is based on self-selection and is not representative of the Swiss population. Subjects are randomly assigned[61] to the different groups in order to create a pre-treatment similarity of the groups with respect to relevant variables, such as gender, age, income, education, and political interest. Thus, possible threats to the effects due to confounding variables are randomly distributed over conditions. Subjects in the different groups tend to have the same average characteristics; the only systematic difference is the treatment (Shadish, Cook, & Campbell, 2002, p. 248ff.). Consequently, changes in the outcome variable are not caused by differences of personal characteristics between the groups (McDermott, 2002a). Thus, random assignment facilitates casual inference.

In the control group (n = 157), 71 percent were males, the age ranged from 19 to 84 (M = 42; SD = 14.5), and 69 percent had a higher education entrance qualification or a higher level of formal education. In the conflict group (n = 189), 67 percent were males, the age ranged from 19 to 76 (M = 44; SD = 15.3), and 68 percent had a higher education entrance qualification or a higher level of formal education. In the inefficiency group (n = 177), 71 percent were males, the age ranged from 18 to 80 (M = 44; SD = 16.1), and 77 percent had a higher education entrance qualification or a higher level of formal education.

6.2.2. Procedures

The study was conceptualized as an internet experiment. Internet experiments are considered to be an efficient way of doing experimental research that makes it more easy to reach diverse populations (Iyengar, 2001). Citizens who sent an e-mail to the

61 Random assignment refers to the fact that units are assigned to conditions based only on chance. Each unit has a nonzero probability of being assigned to a condition.

author and expressed their interest in participating in the study were invited to participate in the research project on May 5, 2008. The initial survey took about 10 minutes to complete. The response rates are 0.82 for the experimental groups and 0.84 for the control group.[62] Rather surprisingly, the higher work load for participants in the experimental groups did not keep subjects from participating in the study. On May 19, 2008 an e-mail invitation to fill out an online questionnaire referring to a news article was sent to the participants in the two treatment groups. From then until May 23, 2008, the subjects in the two treatment groups received, each day, an invitation to an online questionnaire that referred to a news article. Each of the five surveys on news article took about five to eight minutes to complete. Response rates were between 0.89 and 0.92 for the conflict group and 0.86 and 0.90 for the efficiency group.[63] These rates are satisfactory; the treatment articles were received by the study's participants. In order to have a more detailed understanding of how the articles were received, I investigated how many articles were read by the participants in average. The majority of subjects (95 percent) answered the questions to all five articles. There is no significant difference between the two groups as regards how many articles were read. In addition, I looked at when the participants in average read the articles. An average of 66 percent of all participants read an article on the same day on which the article was sent, 18 percent read an article one or two days later, and 11 percent read an article within three to five days later. There is no significant difference between the two groups as regards when the articles were read.

On May 26, 2008 an invitation to participate in the final survey was sent out both to participants in the treatment groups and participants in the control group. The final survey took about 15 minutes to complete. Response rates are .89 for the treatment groups, .92 for the control group. Directly after the participants had completed the final survey, they were transferred to a website that contained the debriefing information.[64]

The quality of causal inferences and generalizations drawn from any research study depends on its measures to enhance validity and rule out alternative explanations. In this study, efforts were made to control for various threats to validity. How-

62 The response rates are calculated based on the complete data set because it is a general characteristic of the survey and not of the data set that is finally used for the analysis. The complete data set does include participants who are younger than 18 years old and participants who do not hold Swiss citizenship. In the complete data set, the number of participants in the conflict group is 209, in the inefficiency group 207, and in the control group 172. The response rate is given for the two experimental groups together, because the random assignment to either the group with conflict articles or the group with inefficiency articles took place after the initial survey was completed.

63 The values for the response rates do not match the attrition rates, because the former is based on the complete data set, whereas the latter is based on the adjusted data set that does not include subjects who are younger than 18 years or do not hold Swiss citizenship.

64 Subjects were informed about the study's interest in media effects on political support. They were also told that two different versions of newspaper articles were presented in the study in order to investigate possible effects. The debriefing also contained information on how to contact the researcher for information on research results.

ever, as there are tradeoffs between the several aspects, there are threats to validity that cannot be controlled with the design. Although most concerns regarding experimental studies refer to a lack in mundane realism (Graber, 2004; McDermott, 2002a), i.e. the similarity between an experimental setting and the real world, they fail the idea of experiments, since

> "the idea of experimentation is not primarily that a specific set of experimental results apply directly to broader conclusions about human behavior in other settings, but rather that these results develop and test theories that then, in turn, and in aggregation, help explain and predict underlying causal mechanisms in more universal issues" (McDermott, 2002a, p. 335f.).

Hence, in experimental designs, internal validity is vital to external validity. Even so, experiments can possess high levels on both dimensions. In this study care is taken to enhance both internal and external validity.

The internal validity of the study refers to the question of whether the relationship between two variables or concepts reflects a causal relationship (Shadish, et al., 2002, p. 53f.). Since the participants are randomly assigned to the two treatment groups and the control group, a *selection bias* – i.e. systematic differences in subjects' characteristics that could cause the effect – is unlikely. Randomization should ensure that the three groups are equal on relevant variables, i.e. the initial levels of political support, process-preferences and process perceptions as well as political interest and socio-demographic variables. A comparison between the two treatment groups and the control groups in terms of the allocation of characteristic variables shows that the randomization was successful.[65] Table 6.1 gives an overview of the results from the mean comparisons with regard to age, income, education, political placement, radio use, regional newspaper use, national newspaper use, tabloid paper use, free newspaper use, television use, and internet use. There is only one significant difference between the groups: The subjects in the inefficiency treatment group read national newspapers less intensely than the participants in the control group and participants in the conflict treatment groups. Additional Chi-Square tests were conducted to test for differences as regards gender, the frequency of direct experiences with politics, and the frequency of indirect experiences with politics. These tests revealed no significant differences between the different conditions on gender[66], the

65 The randomization analysis is based on the final data set in order to test whether the randomization works for those participants who are actually included in the data analysis of media effects. The final data set does not include participants who are younger than 18 years old and participants that do not hold Swiss citizenship. Moreover, all subjects who participated in the initial survey only were excluded from the final data set (for more information on data preparation, see Section 5.3.3). In the final data set the number of participants in the conflict group is 189, in the inefficiency group 177, and in the control group 157.

66 69 percent men in the conflict group, 71 percent men in the inefficiency group, 71 percent men in the control group, Chi-Square = .931, df = 2, p =.628

frequency of direct experiences with politics[67], and the frequency with indirect experiences with politics.[68]

History is a threat to internal validity in the sense that events occurring concurrently with the treatment could also cause the effect. However, the design controls for history, as participants in the control group and the treatment groups are exposed to the same external events. Hence, any possible effects on political support of such external events (i.e. political scandals or economic crisis) have an equal effect on the levels of support of participants in all groups. Moreover, in order to minimize the possibility of impacts of external effects associated with election campaigns, a period of regular political decision-making was chosen as the timeframe for the experimental study. The study was conducted during summer 2008 as a non-election period.

In order to avoid unwanted *sensitization effects*, no pretest measurements on the dependent variables were included in the initial survey. Nosanchuck & Marchak (1969) reported that sensitization is the main unwanted effect of pretests. Pretests may draw the attention to the purpose of the study, which results in increased attention to relevant aspects of the study. This has consequences for the outcome variables. For instance, Nosanchuck & Marchak (1969) found consistency effects[69] when a pretest was used. As a consequence, subsequent attitude changes might be suppressed. Based on a review of experimental studies that include pretests, Bracht & Glass (1968) concluded that sensitization to subsequent treatment is most likely to occur when the dependent variable is a self-report measure. As in the present study the dependent variable is a self-reported measure, measurements of political support, process preferences, and process perceptions were not included in the initial survey that takes place before the treatment is implemented. Thus, the questionnaire in the initial survey focuses on socio-demographic variables and is kept rather short. This also contributes to the prevention of subject fatigue due to long item batteries. Besides the advantages associated with the lack of a pretest, there are also drawbacks. The lack of a pretest is particularly risky, if there is a likelihood of attrition from the study. Without pretest measures, it is more difficult to determine whether those subjects who had dropped out of the study differ from the participants who continued with respect to their characteristics on the dependent variables. Hence, in order to compensate for the abdication of a pretest, several measures to minimize the risk of attrition were taken.

67 38 percent no direct experiences in the conflict group, 34 percent no direct experiences in the inefficiency group, 37 percent no direct experiences in the control group, Chi-Square = 1.464, df = 4, p = .833

68 34 percent no indirect experiences in the conflict group, 34 percent no indirect experiences in the inefficiency group, 38 percent no indirect experiences in the control group, Chi-Square = 1.675, df = 4, p =. 795

69 Consistency effects refer to the desire of the respondents to appear consistent in answering what they perceive to be related questions. As a consequence, the answer to one question is moved in the direction to the other question (Weisberg, 2005, p. 119f.; Weisberg, Krosnick, & Bowen, 1996, p. 97).

Variable	Mean			SD			F	p
	Conflict Group	Inefficiency Group	Control Group	Conflict Group	Inefficiency Group	Control Group		
Age (in years)	43.6	44.48	42.23	15.32	16.13	14.52	0.88	.41
Income (16 = 15.000 CHF and more)	8.37	8.73	8.74	3.74	3.6	3.97	0.53	.89
Education (6 = college or university degree)	4.9	5.04	5.2	1.19	1.27	1.22	2.68	.07
Political placemen(11 = far right)	4.94	5.16	5.02	2.18	2.33	2.09	0.38	.68
Political interest (4 = very intcrcstcd)	3.56	3.49	3.56	0.53	0.54	0.58	0.86	.43
Radio use (4 = more than 1 hour)	1.8	1.56	1.73	1.09	1.3	1.23	1.84	.16
Regional newspaper use (4 = more than 1 hour)	1.16	1.23	1.31	1.05	1.16	1.08	0.77	.47
National newspaper use (4 = more than 1 hour)	1.92	1.51	1.88	1.22	1.25	1.23	5.11	.01
Boulevard paper use (4 = more than 1 hour)	0.18	0.15	0.17	0.5	0.41	0.38	0.11	.90
Free newspaper use (4 = more than 1 hour)	0.68	0.63	0.75	0.79	0.65	0.69	1.07	.34
Television use (4 = more than 1 hour)	1.96	1.74	1.82	1.28	1.29	1.21	1.87	.31
Internet use (4 = more than 1 hour)	1.35	1.36	1.55	1.15	1.1	1.05	1.75	.18

Note. Entries are the Standardized Means, Standard Deviations, F-Values, and significance of propability (two-tailed, p-value), with N between 132 (boulevard use in control group) and 188 (political interest in conflict group)

Table 6.1. Results of Randomization Check

Whereas several threats to internal validity are controlled for with the proposed design, *attrition*[70] remains an uncontrolled threat to internal validity. Attrition refers to the loss of responses from participants who drop out of the study after the subjects were randomly assigned to conditions (Shadish, et al., 2002, p. 323). Attrition can happen before the treatment is implemented or after the treatment is implemented. Post-treatment attrition is a subset of selection bias that occurs after the treatment took place (Shadish, et al., 2002, p. 59). Attrition is a problem if the participants who drop out of the study differ from those who continue to participate. Differences in the outcome measure could then be caused by that loss rather than by the treatment. In the present study, it will not be possible to prevent any attrition given the rather long duration of the experimental study. Although attrition cannot be controlled for, several measures were used to prevent attrition (cf. Shadish, et al., 2002, pp. 323-340). Questionnaires were kept as short as possible. In addition, the initial invitation to participate in the study contained detailed information about the structure of the study (number of questionnaires, approximate time it takes to complete the questionnaires, information about when the questionnaires will be sent) as well as the nature of the media stimuli, in order to avoid surprises for the participants later on. Subjects who did not participate in a survey were reminded to please do so in order to increase response rates and prevent attrition. Because research indicates that attrition is lower, when the time between random assignment and treatment implementation is minimized (Shadish, et al., 2002, p. 331), the time span between the initial survey and the treatment phase was kept as short as possible. In that way, the time span should still allow enough time for the participants to complete the initial survey. Thus, the treatment was implemented two weeks after the invitation to participate in the initial survey had been sent. Most importantly, incentives were offered as compensation for the study participants' time and in order to motivate participants to stay in the study. There is empirical evidence that incentives are useful and have a positive effect on response rate (Harkness, Mohler, Schneid, & Christoph, 1998). All participants in the two treatment groups who completed the experimental study (initial and final survey plus at least one questionnaire on news article) could choose between a voucher for books to a value of 20 CHF and a voucher for the cinema (value of 18 CHF). Subjects in the control group who completed the initial and the final survey received a book voucher to the value of 10 CHF.

To be able to analyze whether people who drop out of the study differed from the ones who do not, relevant variables were measured in the initial survey conducted before the treatment takes place. This information enables attrition to be analyzed, in order to understand how much it threatens the validity of a conclusion about the treatment effectiveness (Shadish, et al., 2002, p. 334ff.).[71] The rates of attrition from

70 Sometimes attrition is also called mortality.

71 The attrition analysis is based on the adjusted data set that does not include participants who are younger than 18 years old and participants who do not hold Swiss citizenship. However, subjects who participated in the initial survey only are part of this adjusted data set because

the initial survey to the final survey are 9.1 percent for the conflict group, 11.5 percent for the inefficiency group, and 8.2 percent for the control group. Comparing the initial survey with the response rate for the five surveys on the news articles, attrition ranges between 4.7 percent and 10.5 percent for the conflict group, and between 7.7 percent and 11 percent for the inefficiency group. There is almost no attrition from the article surveys to the final survey, indicating that those subjects who had the treatment also participated in the posttests. Hence, most attrition occurred before the treatment was implemented. Generally, if attrition is low (less than 10 percent) it can be neglected (Shadish, et al., 2002, p. 229f.). Considering the sample of participants in the conflict group, the inefficiency group and the control group together, there are no significant differences between subjects who continued to participate and those who dropped out on relevant variables, e.g. variables that might be correlated with the outcome (age, income, education, political placement, political interest, media use; for all variables: $df = 5$, $F < 2.2$). There are also no differences between dropouts and individuals who continued participation as regards gender (Chi-Square = 1.304, $df = 5$, $p = .935$), direct experiences with politics (Chi-Square = 15.886, $df = 10$, $p=.103$), and indirect experiences with politics (Chi-Square = 9.856, $df = 10$, $p = .453$). In general, then, attrition appears to be no threat to internal validity in this study.

Regarding uncontrolled threats to construct validity of cause, *hypotheses guessing* by the participants as well as *reactivity to the experimental situation* and consequential influences on their responses are considered to be a serious threat. Literature suggests several ways to reduce this problem (cf. Shadish, et al., 2002, p. 77f.). One possible way is using ethically justifiable deceptions to ensure that subjects cannot guess the hypotheses (McDermott, 2002a, p. 41). In order to distract participants from the purpose of this study, they were told that the study is interested in the perception of political news coverage. Hence, questions on how the participants perceive and evaluate the news articles were included in the surveys. In addition, ensuring anonymity and confidentiality can reduce the risk of probands' evaluation apprehensions.

Participants received the survey questionnaires and media stimuli at home or in their office, depending on where they have internet access. Given the field setting, special care must be given to *treatment implementation*, treatment adherence in particular (cf. Shadish, et al., 2002, pp. 316-320; Yeaton & Sechrest, 1981). Treatment adherence in the study's context refers to the question of whether people read the newspaper articles as intended to. In order to improve adherence, the surveys include questions that refer to the stimulus articles. Manipulation checks were also looked at for information on whether treatment was received or not (Shadish, et al., 2002, p. 317). What this study cannot control, however, is whether those subjects to

the analysis of their loss is precisely what attrition measures (for more information on data preparation see Section 7.2.2). In the adjusted data set, the number of participants in the conflict group is 205, in the inefficiency group 200, and in the control group 171.

which the articles and questionnaires were sent to are actually the ones who participated in the study.

6.2.3. Stimulus Material

The *strength, integrity, and effectiveness of a treatment* are central aspects of any experimental study (Yeaton & Sechrest, 1981, p. 156). Treatment strength is defined as the "a priori likelihood that the treatment could have its intended outcome" (Yeaton & Sechrest, 1981, p. 156). Treatment dose, frequency and length of the treatment define its strength. A strong treatment is aimed, because the effect size is considered to be relatively low, given that political support is a rather stable attitude. Regarding the strength of the treatment, different treatment strengths are used in experimental research that explores the effects of media information on political support. Strength ranges from one or two single articles at one point in time (cf. for instance De Vreese, 2004; Bertram Scheufele, 2008; Valentino, Beckmann, et al., 2001) to five days of treatment (cf. Cappella & Jamieson, 1997). In order to ensure treatment strength in the present study, the treatment could be measured over the period of two weeks, which would allow two posttests, one after each week. However, such a strong treatment involves the risk that the participants in the study get tired (subject fatigue) and drop out (attrition). In order to use a treatment that is both intense enough to evoke possible effects on the one hand, but is not too much of a burden for the study's participants on the other hand, the treatment consists of one media article per weekday and lasts over five days. Hence, five articles per participant constitute the treatment. The articles were sent to the respondents as part of online surveys. More precisely, respondents received one survey each day over the period of five consecutive days.

As there is a trade-off between treatment strength and generalizability of cause to real world conditions, efforts will be made to ensure that treatments are both strong and realistic. The authenticity of media stimuli is a very important aspect that shapes the external validity of experimental research (Matthes, 2007b, p. 306f.; Trepte & Wirth, 2004). Thus, this study makes efforts to ensure that the media stimuli are as authentic as possible. A precondition of external validity is that the stimulus material occurs in a similar form in real world situations. One way to ensure the external validity of media stimuli is to investigate regular and characteristic patterns of mass media content and to develop the media stimuli according to these patterns (for instance Cappella & Jamieson, 1996; Gilliam & Iyengar, 2000; Iyengar, 1987). Therefore, results from the content analysis (Chapter 4) are used to inform the development of the media articles that serve as treatment in the experimental study.[72]

72 An alternative procedure could have been to conduct a qualitative content analysis. However, this would not allow us to derive characteristic patterns of media presentations of political processes. Such typical patterns, however, are reflected in media stimuli, which is why the development is based on findings from this quantitative analysis.