

1 A tutorial on Bayesian and Frequentist Event History Analyses for psychological
2 time-to-event data

3 Sven Panis¹ & Richard Ramsey¹

4 ¹ ETH Zürich

5 Author Note

6 Neural Control of Movement lab, Department of Health Sciences and Technology

7 (D-HEST).

8 The authors made the following contributions. Sven Panis: Conceptualization,
9 Writing - Original Draft Preparation, Writing - Review & Editing; Richard Ramsey:
10 Conceptualization, Writing - Review & Editing, Supervision.

11 Correspondence concerning this article should be addressed to Sven Panis, ETH
12 GLC, room G16.2, Gloriastrasse 37/39, 8006 Zürich. E-mail: sven.panis@hest.ethz.ch

13

Abstract

14 Time-to-event data such as response times, saccade latencies, and fixation durations are
15 ubiquitous in experimental psychology. The orthodox method for analysing such data -
16 comparing means with analysis-of-variance - is actually hiding a lot of information about
17 psychological effects, such as their onset time and duration, and whether they are
18 time-locked to stimuli. Such information can change key conclusions about psychological
19 processes and can be revealed by using distributional measures.

20 Here we provide a set of tutorials on how to implement one particular distributional
21 method known as discrete-time event history analysis, a.k.a. hazard analysis, duration
22 analysis, failure-time analysis, survival analysis, and transition analysis. We illustrate how
23 one can calculate the descriptive statistics, and how one can implement Bayesian and
24 frequentist regression models, using the R packages tidyverse, brms, and lme4. The R code
25 is publicly available on Github and OSF, and can easily be adapted for other data sets. We
26 further discuss how to manage inter-individual differences, implications for experimental
27 design, and how to select among various options when analysing time-to-event data using
28 discrete-time hazard analysis.

29 *Keywords:* response times, event history analysis, Bayesian regression models

30 Word count: X

31 A tutorial on Bayesian and Frequentist Event History Analyses for psychological
 32 time-to-event data

33 **Introduction**

34 **Means versus distributional shapes**

35 In experimental psychology, it is still standard practice to analyse response times
 36 (RTs), saccade latencies, and fixation durations by calculating average performance across
 37 a series of trials. However, differences in means conceal when an experimental effect starts,
 38 how long it lasts, how it evolves over increasing waiting time, and whether its onset is
 39 time-locked to other events. Such information is useful not only for interpretation, but also
 40 for cognitive psychophysiology and computational model selection (Panis, Schmidt,
 41 Wolkersdorfer, & Schmidt, 2020). As a simple illustration, Figure 1 shows three examples
 42 of how an observed difference in mean response times (RTs) between two experimental
 43 conditions conceals differences in the shapes of the underlying RT distributions.

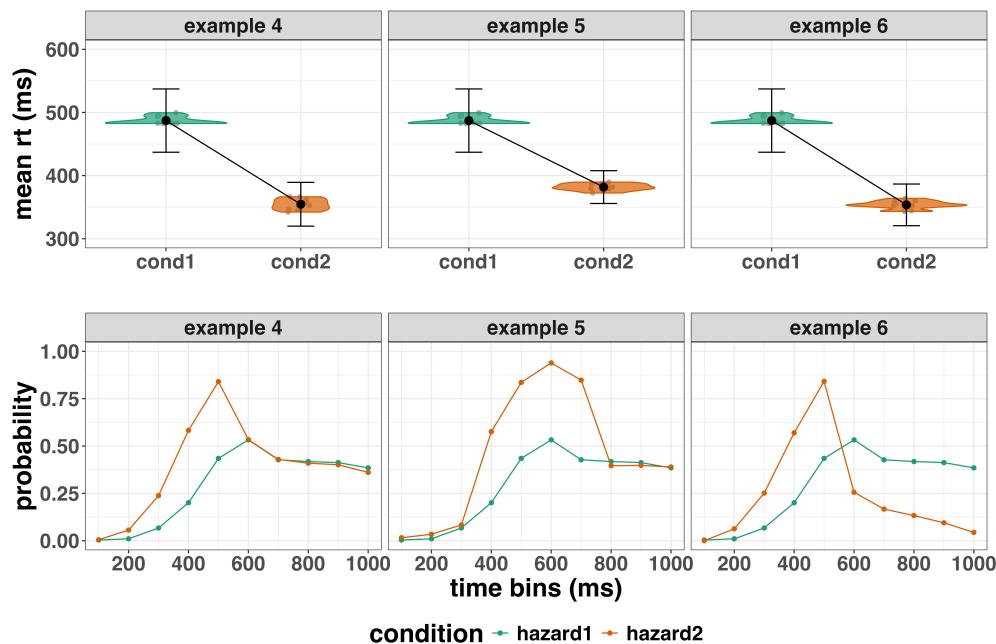


Figure 1. Means versus distributional shapes.

⁴⁴ In each example, the mean RT is lower in condition 2 compared to condition 1.
⁴⁵ However, the distributions in the first example show that the effect starts around 200 ms
⁴⁶ and is gone by 600 ms. In the second example, the distributional effect starts around 400
⁴⁷ ms and is gone by 800 ms. And in the third example, the distributional effect reverses
⁴⁸ around 550 ms.

⁴⁹ **Outline of the paper**

⁵⁰ In this paper we focus on a particular distributional method known as discrete-time
⁵¹ event history analysis, a.k.a. hazard analysis, duration analysis, failure-time analysis,
⁵² survival analysis, and transition analysis. We first provide a brief overview of discrete-time
⁵³ hazard analysis to orient the reader to the basic concepts and ideas that we will use
⁵⁴ throughout the paper. However, this will remain relatively short, and for detailed
⁵⁵ treatment, see Singer and Willett (2003), Allison (1982), and Allison (2010).

⁵⁶ We then provide four different tutorials, each of which is written in R code and
⁵⁷ publicly available on Github and the Open Science Framework (OSF). The tutorials
⁵⁸ provide hands-on, concrete examples of key parts of the analytical process, so that others
⁵⁹ can apply the analyses to their own time-to-event data sets. In Tutorial 1 we illustrate how
⁶⁰ to calculate the descriptive statistics for a published data set when there is one
⁶¹ independent variable. The descriptive statistics are plotted, and we comment on their
⁶² interpretation. In Tutorial 2 we illustrate how one can fit Bayesian hazard models to the
⁶³ data. After selecting the best of four models, we plot the model-based effects and the
⁶⁴ model fits for a few subjects. In Tutorial 3 we illustrate how to fit hazard models in a
⁶⁵ frequentist framework. We compare the model-based effects between Bayesian and
⁶⁶ frequentist approaches. In Tutorial 4 we illustrate how to calculate the descriptive
⁶⁷ statistics when there are two independent variables.

68 Overview of hazard analysis

69 To apply event history analysis (EHA), one must be able to define the event of
70 interest (any qualitative change that can be situated in time, e.g., a button press, a saccade
71 onset, a fixation offset, etc.), time point zero (e.g., target stimulus onset, fixation onset),
72 and measure the passage of time between time point zero and event occurrence in discrete
73 or continuous time units.

74 The shape of a distribution of waiting times can be described in multiple ways (Luce,
75 1991). Let RT be a continuous random variable denoting a particular person's response time
76 in a particular experimental condition. Because waiting times can only increase,
77 continuous-time EHA does not focus on the cumulative distribution function $F(t) = P(RT$
78 $\leq t)$ and its derivative, the probability density function $f(t) = F(t)'$, but on the survivor
79 function $S(t) = P(RT > t)$ and the hazard rate function $\lambda(t) = f(t)/S(t)$. The hazard rate
80 function gives you the instantaneous rate of event occurrence at time point t , given that
81 the event has not occurred yet.

82 Similarly, after dividing time in discrete, contiguous time bins indexed by t , let RT be
83 a discrete random variable denoting the rank of the time bin in which a particular person's
84 response occurs in a particular experimental condition. Discrete-time EHA focuses on the
85 discrete-time hazard function $h(t) = P(RT = t | RT \geq t)$ and the discrete-time survivor
86 function $S(t) = P(RT > t) = [1-h(t)].[1-h(t-1)].[1-h(t-2)] \dots [1-h(1)]$, and not on the
87 probability mass function $p(t) = h(t).S(t-1)$ and the cumulative distribution function $F(t)$
88 = $1-S(t)$. The discrete-time hazard probability function gives you the probability that the
89 event occurs (sometime) in bin t , given that the event has not occurred yet in previous
90 bins. Unlike the hazard function, which assesses the unique risk associated with each time
91 bin, the survivor function cumulates the bin-by-bin risks of event *nonoccurrence*.

92 For two-choice RT data, the discrete-time hazard function can be extended with the
93 conditional accuracy function $ca(t) = P(\text{correct} | RT = t)$, which gives you the probability

94 that a response is correct given that it has been emitted in time bin t (Allison, 2010;
95 Kantowitz & Pachella, 2021; Wickelgren, 1977). This latter function is also known as the
96 micro-level speed-accuracy tradeoff function.

97 Statisticians and mathematical psychologists recommend focusing on the hazard
98 function when analyzing time-to-event data for various reasons. First, as discussed by
99 Holden, Van Orden, and Turvey (2009), “probability density functions can appear nearly
100 identical, both statistically and to the naked eye, and yet are clearly different on the basis
101 of their hazard functions (but not vice versa). Hazard functions are thus more diagnostic
102 than density functions” (p. 331). Second, because RT distributions may differ from one
103 another in multiple ways, Townsend (1990) developed a dominance hierarchy of statistical
104 differences between two arbitrary distributions A and B. For example, if $F_A(t) > F_B(t)$ for
105 all t , then both cumulative distribution functions are said to show a complete ordering.

106 Townsend (1990) showed that a complete ordering on the hazard functions — $\lambda_A(t) > \lambda_B(t)$
107 for all t — implies a complete ordering on both the cumulative distribution and survivor
108 functions — $F_A(t) > F_B(t)$ and $S_A(t) < S_B(t)$ — which in turn implies an ordering on the
109 mean latencies —mean A < mean B. In contrast, an ordering on two means does not imply
110 a complete ordering on the corresponding $F(t)$ and $S(t)$ functions, and a complete ordering
111 on these latter functions does not imply a complete ordering on the corresponding hazard
112 functions. This means that stronger conclusions can be drawn from data when comparing
113 the hazard functions using EHA. For example, when mean A < mean B, the hazard
114 functions might show a complete ordering (i.e., for all t), a partial ordering (e.g., only for t
115 > 300 ms, or only for $t < 500$ ms), or they may cross each other one or more times. Third,
116 EHA does not discard right-censored observations when estimating hazard functions, that
117 is, trials for which we do not observe a response during the data collection period so that
118 we only know that the RT must be larger than some value. This is important because
119 although a few right-censored observations are inevitable in most RT tasks, a lot of
120 right-censored observations are expected in experiments on masking, the attentional blink,

and so forth. Fourth, hazard modeling allows incorporating time-varying explanatory covariates such as heart rate, electroencephalogram (EEG) signal amplitude, gaze location, etc. (Allison, 2010) which is useful for cognitive psychophysiology (Meyer, Osman, Irwin, & Yantis, 1988). Finally, as explained by Kelso, Dumas, and Tognoli (2013), it is crucial to first have a precise description of the macroscopic behavior of a system (here: $h(t)$ and $ca(t)$ functions) in order to know what to derive on the microscopic level. For example, fitting parametric functions or computational models to data without studying the shape of the $h(t)$ and $ca(t)$ functions can miss important features in the data (Panis, Moran, Wolkersdorfer, & Schmidt, 2020; Panis & Schmidt, 2016).

We focus on factorial within-subject designs in which a large number of observations are made on a relatively small number of participants (small- N designs). This approach emphasizes the precision and reproducibility of data patterns at the individual participant level to increase the inferential validity of the design (Baker et al., 2021; Smith & Little, 2018). In contrast to the large- N design that averages across many participants without being able to scrutinize individual data patterns, small- N designs retain crucial information about the data patterns of individual observers. This is of great advantage whenever participants differ systematically in their strategies or in the time-courses of their effects, so that blindly averaging them would lead to misleading data patterns. Indeed, Smith and Little (2018) argue that, “if psychology is to be a mature quantitative science, then its primary theoretical aim should be to investigate systematic functional relationships as they are manifested at the individual participant level” (p. 2083). Note that because statistical power derives both from the number of participants and from the number of repeated measures per participant and condition, small- N designs can have excellent power (Baker et al., 2021; Smith & Little, 2018).

We used R (Version 4.4.0; R Core Team, 2024)¹ for all reported analyses. Web links

¹ We, furthermore, used the R-packages *bayesplot* (Version 1.11.1; Gabry, Simpson, Vehtari, Betancourt, & Gelman, 2019), *brms* (Version 2.21.0; Bürkner, 2017, 2018, 2021), *citr* (Version 0.3.2; Aust, 2019),

¹⁴⁶ are printed in bold. The content of the tutorials is mainly based on Allison (2010), Singer
¹⁴⁷ and Willett (2003), McElreath (2018), Kurz (2023a), and Kurz (2023b).

¹⁴⁸

Tutorial 1: Calculating descriptive statistics using a life table

¹⁴⁹ To illustrate how to quickly set up life tables for calculating the descriptive statistics
¹⁵⁰ (functions of discrete time), we use a published data set on masked response priming from
¹⁵¹ Panis and Schmidt (2016), available on **ResearchGate**. In their first experiment, Panis
¹⁵² and Schmidt (2016) presented a double arrow for 94 ms that pointed left or right as the
¹⁵³ target stimulus with an onset at time point zero in each trial. Participants had to indicate
¹⁵⁴ the direction in which the double arrow pointed using their corresponding index finger,
¹⁵⁵ within 800 ms after target onset. Response time and accuracy were recorded on each trial.
¹⁵⁶ Prime type (blank, congruent, incongruent) and mask type were manipulated. Here we
¹⁵⁷ focus on the subset of trials in which no mask was presented. The 13-ms prime stimulus
¹⁵⁸ was a double arrow with onset at -187 ms for the congruent (same direction as target) and
¹⁵⁹ incongruent (opposite direction as target) prime conditions.

cmdstanr (Version 0.8.1.9000; Gabry, Češnovar, Johnson, & Bronder, 2024), *dplyr* (Version 1.1.4;
Wickham, François, Henry, Müller, & Vaughan, 2023), *forcats* (Version 1.0.0; Wickham, 2023a), *ggplot2*
(Version 3.5.1; Wickham, 2016), *lme4* (Version 1.1.35.5; Bates, Mächler, Bolker, & Walker, 2015), *lubridate*
(Version 1.9.3; Grolemund & Wickham, 2011), *Matrix* (Version 1.7.0; Bates, Maechler, & Jagan, 2024),
nlme (Version 3.1.166; Pinheiro & Bates, 2000), *papaja* (Version 0.1.2.9000; Aust & Barth, 2023),
patchwork (Version 1.2.0; Pedersen, 2024), *purrr* (Version 1.0.2; Wickham & Henry, 2023), *RColorBrewer*
(Version 1.1.3; Neuwirth, 2022), *Rcpp* (Eddelbuettel & Balamuta, 2018; Version 1.0.12; Eddelbuettel &
François, 2011), *readr* (Version 2.1.5; Wickham, Hester, & Bryan, 2024), *RJ-2021-048* (Bengtsson, 2021,
2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021,
2021, 2021, 2021, 2021), *standist* (Version 0.0.0.9000; Girard, 2024), *stringr* (Version 1.5.1; Wickham,
2023b), *tibble* (Version 3.2.1; Müller & Wickham, 2023), *tidybayes* (Version 3.0.6; Kay, 2023), *tidyverse* (Version 1.3.1;
Wickham, Vaughan, & Girlich, 2024), *tidyverse* (Version 2.0.0; Wickham et al., 2019), and *tinylabels*
(Version 0.2.4; Barth, 2023).

160 After loading in the data file, one has to (a) supply required column names, and (b)
 161 specify the factor condition with the correct levels and labels. The required column names
 162 are as follows:

- 163 • “pid”, indicating unique participant IDs;
- 164 • “trial”, indicating each unique trial per participant;
- 165 • “condition”, a factor indicating the levels of the independent variable (1, 2, ...) and
 166 the corresponding labels;
- 167 • “rt”, indicating the response times in ms;
- 168 • “acc”, indicating the accuracies (1/0).

169 In the code of Tutorial 1, this is accomplished as follows.

```
data_wr <- read_csv("../Tutorial_1_descriptive_stats/data/DataExp1_6subjects_wrangled.csv")
colnames(data_wr) <- c("pid", "bl", "tr", "condition", "resp", "acc", "rt", "trial")
data_wr <- data_wr %>%
  mutate(condition = condition + 1, # original levels were 0, 1, 2.
        condition = factor(condition, levels=c(1,2,3), labels=c("blank", "congruent", "incongruent")))

```

170 To set up the life tables and plots of the discrete-time functions $h(t)$, $S(t)$ and $ca(t)$
 171 using functional programming, one has to nest the data within participants using the
 172 `group_nest()` function, and supply a user-defined censoring time and bin width to our
 173 function “`censor()`”, as follows.

```
data_nested <- data_wr %>% group_nest(pid)
data_final <- data_nested %>%
  mutate(censored = map(data, censor, 600, 40)) %>% # ! user input: censoring time, and bin width
  mutate(ptb_data = map(censored, ptb)) %>% # create person-trial-bin dataset
  mutate(lifetable = map(ptb_data, setup_lt)) %>% # create life tables without ca(t)
  mutate(condacc = map(censored, calc_ca)) %>% # calculate ca(t)
  mutate(lifetable_ca = map2(lifetable, condacc, join_lt_ca)) %>% # create life tables with ca(t)
  mutate(plot = map2(.x = lifetable_ca, .y = pid, plot_eha, 1)) # create plots
```

174 Note that the censoring time should be a multiple of the bin width (both in ms). The
 175 censoring time should be a time point after which no informative responses are expected

176 anymore. In experiments that implement a response deadline in each trial the censoring
177 time can equal that deadline time point. Trials with a RT larger than the censoring time,
178 or trials in which no response is emitted during the data collection period, are treated as
179 right-censored observations in EHA. In other words, these trials are not discarded, because
180 they contain the information that the event did not occur before the censoring time.

181 Removing such trials before calculating the mean event time can introduce a sampling bias.
182 The person-trial-bin oriented dataset has one row for each time bin of each trial that is at
183 risk for event occurrence. The variable “event” in the person-trial-bin oriented data set
184 indicates whether a response occurs (1) or not (0) for each bin. When creating the plots
185 using our function `plot_eha()`, some warning messages will likely be generated, like these:

- 186 • Removed 2 rows containing missing values or values outside the scale range
187 (`geom_line()`).
- 188 • Removed 2 rows containing missing values or values outside the scale range
189 (`geom_point()`).
- 190 • Removed 2 rows containing missing values or values outside the scale range
191 (`geom_segment()`).

192 The warning messages are generated because some bins have no hazard and $ca(t)$
193 estimates, and no error bars. They can thus safely be ignored. One can now inspect
194 different aspects, including the life table for a particular condition of a particular subject,
195 and a plot of the different functions for a particular participant.

196 Table 1 shows the life table for condition “blank” (no prime stimulus presented) -
197 compare to Figure 1. A life table includes for each time bin, the risk set (number of trials
198 that are event-free at the start of the bin), the number of observed events, and the
199 estimates of $h(t)$, $S(t)$, $ca(t)$ and their estimated standard errors (se). At time point zero,
200 no events can occur and therefore $h(t)$ and $ca(t)$ are undefined.

201 Figure 1 displays the discrete-time hazard, survivor, and conditional accuracy
202 functions for each prime condition for participant 6. By using discrete-time $h(t)$ functions
203 of event occurrence - in combination with $ca(t)$ functions for two-choice tasks - one can
204 provide an unbiased, time-varying, and probabilistic description of the latency and
205 accuracy of responses based on all trials of any data set.

206 For example, for participant 6, the estimated hazard values in bin (240,280] are 0.03,
207 0.17, and 0.20 for the blank, congruent, and incongruent prime conditions, respectively. In
208 other words, when the waiting time has increased until *240 ms* after target onset, then the
209 conditional probability of response occurrence in the next 40 ms is more than five times
210 larger for both prime-present conditions, compared to the blank prime condition.

211 Furthermore, the estimated conditional accuracy values in bin (240,280] are 0.29, 1,
212 and 0 for the blank, congruent, and incongruent prime conditions, respectively. In other
213 words, if a response is emitted in bin (240,280], then the probability that it is correct is
214 estimated to be 0.29, 1, and 0 for the blank, congruent, and incongruent prime conditions,
215 respectively.

Table 1

The life table for the blank prime condition of participant 6.

bin	risk_set	events	hazard	se_haz	survival	se_surv	ca	se_ca
0.00	220.00	NA	NA	NA	1.00	0.00	NA	NA
40.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
80.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
120.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
160.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
200.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
240.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
280.00	220.00	7.00	0.03	0.01	0.97	0.01	0.29	0.17
320.00	213.00	13.00	0.06	0.02	0.91	0.02	0.77	0.12
360.00	200.00	26.00	0.13	0.02	0.79	0.03	0.92	0.05
400.00	174.00	40.00	0.23	0.03	0.61	0.03	1.00	0.00
440.00	134.00	48.00	0.36	0.04	0.39	0.03	0.98	0.02
480.00	86.00	37.00	0.43	0.05	0.22	0.03	1.00	0.00
520.00	49.00	32.00	0.65	0.07	0.08	0.02	1.00	0.00
560.00	17.00	9.00	0.53	0.12	0.04	0.01	1.00	0.00
600.00	8.00	4.00	0.50	0.18	0.02	0.01	1.00	0.00

Note. The column named “bin” indicates the endpoint of each time bin (in ms), and includes time point zero. For example the first bin is (0,40] with the starting point excluded and the endpoint included. se = standard error. ca = conditional accuracy.

Subject 6

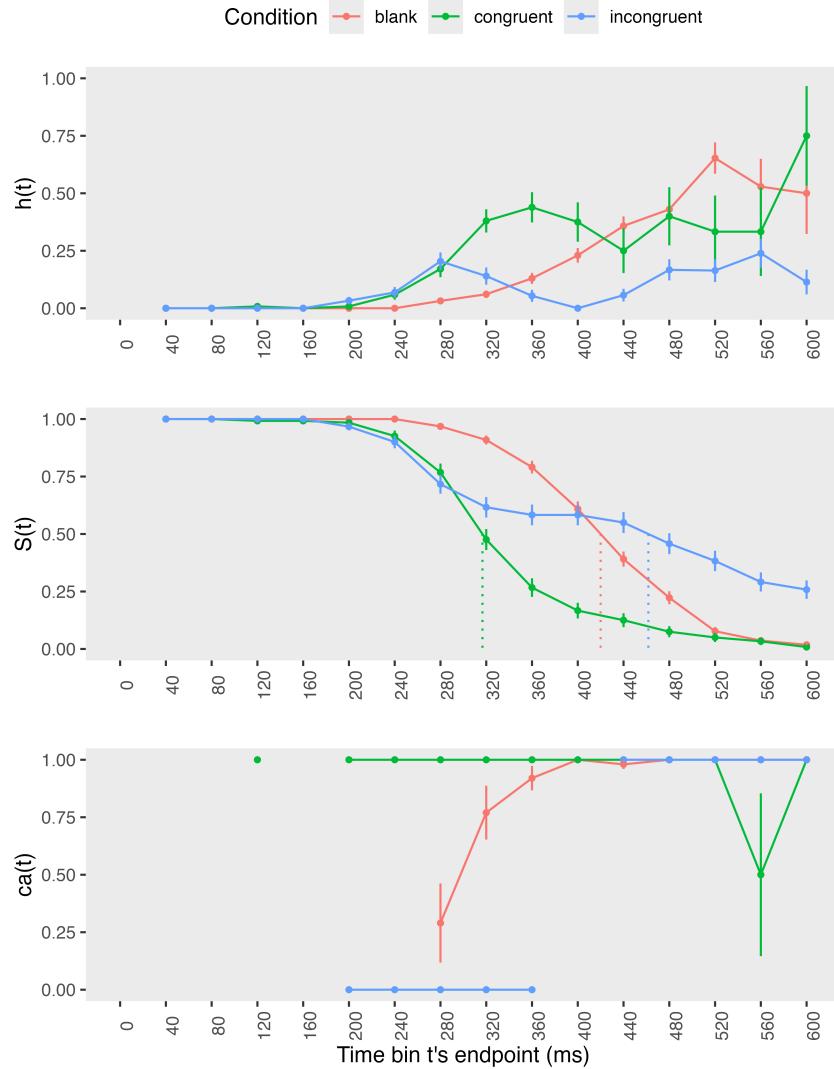


Figure 2. Estimated discrete-time hazard, survivor, and conditional accuracy functions for participant 6, as a function of the passage of discrete waiting time.

216 However, when the waiting time has increased until 400 ms after target onset, then
 217 the conditional probability of response occurrence in the next 40 ms is estimated to be
 218 0.36, 0.25, and 0.06 for the blank, congruent, and incongruent prime conditions,
 219 respectively. And when a response does occur in bin (400,440], then the probability that it
 220 is correct is estimated to be 0.98, 1, and 1 for the blank, congruent, and incongruent prime

221 conditions, respectively.

222 These results show that this participant is initially responding to the prime even
223 though (s)he was instructed to only respond to the target, that response competition
224 emerges in the incongruent prime condition around 300 ms, and that only later response
225 are fully controlled by the target stimulus. Qualitatively similar results were obtained for
226 the other five participants. Also, in their second Experiment, Panis and Schmidt (2016)
227 showed that the negative compatibility effect in the mask-present conditions is time-locked
228 to mask onset. This example shows that a simple difference between two means fails to
229 reveal the dynamic behavior people display in many experimental paradigms (Panis, 2020;
230 Panis, Moran, et al., 2020; Panis & Schmidt, 2022; Panis, Torfs, Gillebert, Wagemans, &
231 Humphreys, 2017; Panis & Wagemans, 2009; Schmidt, Panis, Wolkersdorfer, & Vorberg,
232 2022). In other words, statistically controlling for the passage of time during data analysis
233 is equally important as experimental control during the design of an experiment, to better
234 understand human behavior in experimental paradigms. As we will show in Tutorials 2 and
235 3, statistical models for $h(t)$ and $ca(t)$ can each be implemented as generalized linear mixed
236 regression models predicting event occurrence (1/0) and response accuracy (1/0) in each
237 bin of a selected time range, respectively.

238 **Tutorial 2: Fitting Bayesian hazard models**

239 When you want to study how hazard depends on various predictors, you can fit
240 regression models to the data (Singer & Willett, 2003). There are two analytic decisions
241 one has to make. First, one has to select an analysis time range, i.e., a contiguous set of
242 bins for which there is enough data for each participant. Second, one can choose the logit
243 link function which transforms a (hazard) probability into the log of the odds ratio, or the
244 complementary log-log (cloglog) link function, which yields the logarithm of the negated
245 logarithm of the probability of event *nonoccurrence*. An important difference between
246 these two link functions is that cloglog provides a discrete-time hazard model that has a

²⁴⁷ built-in proportional hazards assumption, while logit provides a proportional odds
²⁴⁸ assumption (see below). The cloglog link is preferred over the logit link when events can
²⁴⁹ occur in principle at any time point within a bin, which is the case for RT data (Singer &
²⁵⁰ Willett, 2003). Third, one has to choose a specification of the effect of discrete TIME (i.e.,
²⁵¹ the time bin index t). One can choose a general specification (one intercept per bin) or a
²⁵² functional specification, such as a polynomial one.

²⁵³ An example discrete-time hazard model with three predictors (TIME, X1, X2) and
²⁵⁴ the cloglog link function can be written as follows:

$$\begin{aligned} \text{cloglog}[h(t)] = \ln(-\ln[1-h(t)]) = [\alpha_1 \text{ONE} + \alpha_2(\text{TIME}-1) + \alpha_3(\text{TIME}-1)^2] + [\beta_1 X_1 + \\ \beta_2 X_2 + \beta_3 X_2(\text{TIME}-1)]. \end{aligned}$$

²⁵⁷ The main predictor variable TIME is the time bin index t that is centered on value 1
²⁵⁸ in this example. The first set of terms within brackets, the alpha parameters multiplied by
²⁵⁹ their polynomial specifications of (centered) time, represents the shape of the baseline
²⁶⁰ cloglog-hazard function (i.e., when all predictors X_i take on a value of zero). The second
²⁶¹ set of terms (the beta parameters) represents the vertical shift in the baseline
²⁶² cloglog-hazard for a 1 unit increase in the respective predictor. Predictors can be discrete,
²⁶³ continuous, and time-varying or time-invariant. For example, the effect of a 1 unit increase
²⁶⁴ in X_1 is to vertically shift the whole baseline cloglog-hazard function by β_1 cloglog-hazard
²⁶⁵ units. However, if the predictor interacts linearly with time (see X_2 in the example), then
²⁶⁶ the effect of a 1 unit increase in X_2 is to vertically shift the predicted cloglog-hazard in bin
²⁶⁷ 1 by β_2 cloglog-hazard units (when $\text{TIME}-1 = 0$), in bin 2 by $\beta_2 + \beta_3$ cloglog-hazard units
²⁶⁸ (when $\text{TIME}-1 = 1$), and so forth. To interpret the effects of the predictors, the parameter
²⁶⁹ estimates are exponentiated, resulting in a hazard ratio (due to the use of the cloglog link).

²⁷⁰ In the case of a large- N design without repeated measurements, the parameters of a
²⁷¹ discrete-time hazard model can be estimated using standard logistic regression software
²⁷² (after expanding the typical person-trial-oriented data set into a person-trial-bin-oriented

273 data set (see Tutorial 1); Allison (2010)). When there is clustering in the data, as in the
 274 case of a small- N design with repeated measurements, the parameters of a discrete-time
 275 hazard model can be estimated using population-averaged methods (e.g., Generalized
 276 Estimating Equations), and Bayesian or frequentist generalized linear mixed models
 277 (Allison, 2010).

278 In this second tutorial we illustrate how to fit a Bayesian hazard regression model for
 279 the masked response priming data set used in the first tutorial. In general, there are three
 280 assumptions one can make or relax when adding experimental predictor variables: The
 281 linearity assumption for continuous predictors (the effect of a 1 unit change is the same
 282 anywhere on the scale), the additivity assumption (predictors do not interact), and the
 283 proportionality assumption (predictors do not interact with TIME).

284 First, we select the analysis range (200,600] and the cloglog link, and use a
 285 polynomial to specify the effect of TIME in the “blank” prime condition. Second, based on
 286 previous work (Panis, 2020; Panis, Moran, et al., 2020; Panis & Schmidt, 2022; Panis et al.,
 287 2017; Panis & Wagemans, 2009) and because cognition is likely the behavior of a non-linear
 288 dynamical system [ref], we relax all three assumptions, as follows:

```
# load data
ptb_data <- read_csv("../Tutorial_1_descriptive_stats/data/inputfile_hazard_modeling.csv")
# select analysis time range: (200,600] with 10 bins (time bin ranks 6 to 15)
ptb_data <- ptb_data %>% filter(period > 5)
# create factor condition, with "blank" as the reference level
ptb_data <- ptb_data %>% mutate(condition = factor(condition, labels = c("blank", "congruent", "incongruent")))
# center discrete TIME (period) on bin 9, and trial on trial 1000
ptb_data <- ptb_data %>% mutate(period_9 = period - 9,
                                 trial_c = (trial - 1000)/1000)
# remove unnecessary columns before fitting a model
ptb_data <- ptb_data %>% select(-c(bl,tr,trial,period)) # 12840 obs. of 5 variables

priors <- c(
  set_prior("normal(0, 1)", class = "b"), # for beta parameters
  set_prior("student_t(7.61, 0, 1.57)", class = "b", coef = "Intercept"), # flat prior for intercept on hazard scale
  set_prior("normal(0, 1)", class = "sd"), # for standard deviation of RE
```

```

set_prior("lkj(2)", class = "cor") # for correlations between RE
)

#plan(multicore)
#model_full_RE <-
#  brm(data = ptb_data,
#       family = binomial(link="cloglog"),
#       event / trials(1) ~ 0 + Intercept +
#                           condition*period_9*trial_c +
#                           condition*I(period_9^2) +
#                           condition*I(period_9^3) +
#                           (1 + condition*period_9*trial_c +
#                           condition*I(period_9^2) +
#                           condition*I(period_9^3) / pid),
#       prior = priors,
#       chains = 4, cores = 4, iter = 3000, warmup = 1000,
#       control = list(adapt_delta = 0.999, step_size = 0.04, max_treedepth = 12),
#       seed = 12, init = "O",
#       file = "../Tutorial_2_Bayesian/models/model_full_RE")

```

289 To test whether (centered) trial number affects behavior, we fit a model without the
 290 variable trial_c.

291 Use WAIC to compare models.

292 Plot the effects of congruent and incongruent for each time bin for the selected model.

293 Plot the model-based hazard and survivor functions.

294 **Tutorial 3: Fitting Frequentist hazard models**

295 In this third tutorial we illustrate how to fit a frequentist hazard regression model for
 296 the data set used in the first tutorial.

Tutorial 4: Calculating descriptive statistics when there are two independent variables

In this final tutorial we illustrate how to calculate and plot the descriptive statistics

³⁰⁰ for the full data set of Experiment 1 of Panis and Schmidt (2016).

Discussion

302 Individual differences

- role of response deadlines, low-level vs. higher-level processes,
 - clustering algorithms based on $h(t)$ and $ca(t)$ data
 -

306 Cognitive psychophysiology and computational model selection

307 Power analysis

- example repo on github

309 Preregistration

- example preregistration for knot data

Conclusions

312

References

- 313 Allison, P. D. (1982). Discrete-Time Methods for the Analysis of Event Histories.
- 314 *Sociological Methodology*, 13, 61. <https://doi.org/10.2307/270718>
- 315 Allison, P. D. (2010). *Survival analysis using SAS: A practical guide* (2. ed). Cary, NC:
- 316 SAS Press.
- 317 Aust, F. (2019). *Citr: 'RStudio' add-in to insert markdown citations*. Retrieved from
- 318 <https://github.com/crsh/citr>
- 319 Aust, F., & Barth, M. (2023). *papaja: Prepare reproducible APA journal articles with R*
- 320 *Markdown*. Retrieved from <https://github.com/crsh/papaja>
- 321 Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., &
- 322 Andrews, T. J. (2021). Power contours: Optimising sample size and precision in
- 323 experimental psychology and human neuroscience. *Psychological Methods*, 26(3),
- 324 295–314. <https://doi.org/10.1037/met0000337>
- 325 Barth, M. (2023). *tinylabes: Lightweight variable labels*. Retrieved from
- 326 <https://cran.r-project.org/package=tinylabes>
- 327 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects
- 328 models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- 329 <https://doi.org/10.18637/jss.v067.i01>
- 330 Bates, D., Maechler, M., & Jagan, M. (2024). *Matrix: Sparse and dense matrix classes and*
- 331 *methods*. Retrieved from <https://CRAN.R-project.org/package=Matrix>
- 332 Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in r
- 333 using futures. *The R Journal*, 13(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>
- 334 Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan.
- 335 *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- 336 Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms.
- 337 *The R Journal*, 10(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>
- 338 Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal*

- 339 *of Statistical Software*, 100(5), 1–54. <https://doi.org/10.18637/jss.v100.i05>
- 340 Eddelbuettel, D., & Balamuta, J. J. (2018). Extending R with C++: A Brief Introduction
341 to Rcpp. *The American Statistician*, 72(1), 28–36.
342 <https://doi.org/10.1080/00031305.2017.1375990>
- 343 Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal
344 of Statistical Software*, 40(8), 1–18. <https://doi.org/10.18637/jss.v040.i08>
- 345 Gabry, J., Češnovar, R., Johnson, A., & Brodner, S. (2024). *Cmdstanr: R interface to
346 'CmdStan'*. Retrieved from <https://github.com/stan-dev/cmdstanr>
- 347 Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization
348 in bayesian workflow. *J. R. Stat. Soc. A*, 182, 389–402.
349 <https://doi.org/10.1111/rssa.12378>
- 350 Girard, J. (2024). *Standist: What the package does (one line, title case)*. Retrieved from
351 <https://github.com/jmgirard/standist>
- 352 Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate.
353 *Journal of Statistical Software*, 40(3), 1–25. Retrieved from
354 <https://www.jstatsoft.org/v40/i03/>
- 355 Holden, J. G., Van Orden, G. C., & Turvey, M. T. (2009). Dispersion of response times
356 reveals cognitive dynamics. *Psychological Review*, 116(2), 318–342.
357 <https://doi.org/10.1037/a0014849>
- 358 Kantowitz, B. H., & Pachella, R. G. (2021). The Interpretation of Reaction Time in
359 Information-Processing Research 1. *Human Information Processing*, 41–82.
360 <https://doi.org/10.4324/9781003176688-2>
- 361 Kay, M. (2023). *tidybayes: Tidy data and geoms for Bayesian models*.
362 <https://doi.org/10.5281/zenodo.1308151>
- 363 Kelso, J. A. S., Dumas, G., & Tognoli, E. (2013). Outline of a general theory of behavior
364 and brain coordination. *Neural Networks: The Official Journal of the International
365 Neural Network Society*, 37, 120–131. <https://doi.org/10.1016/j.neunet.2012.09.003>

- 366 Kurz, A. S. (2023a). *Applied longitudinal data analysis in brms and the tidyverse* (version
367 0.0.3). Retrieved from <https://bookdown.org/content/4253/>
- 368 Kurz, A. S. (2023b). *Statistical rethinking with brms, ggplot2, and the tidyverse: Second*
369 *edition* (version 0.4.0). Retrieved from <https://bookdown.org/content/4857/>
- 370 Luce, R. D. (1991). *Response times: Their role in inferring elementary mental organization*
371 (1. issued as paperback). Oxford: Univ. Press.
- 372 McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and*
373 *Stan* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315372495>
- 374 Meyer, D. E., Osman, A. M., Irwin, D. E., & Yantis, S. (1988). Modern mental
375 chronometry. *Biological Psychology*, 26(1-3), 3–67.
376 [https://doi.org/10.1016/0301-0511\(88\)90013-0](https://doi.org/10.1016/0301-0511(88)90013-0)
- 377 Müller, K., & Wickham, H. (2023). *Tibble: Simple data frames*. Retrieved from
378 <https://CRAN.R-project.org/package=tibble>
- 379 Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes*. Retrieved from
380 <https://CRAN.R-project.org/package=RColorBrewer>
- 381 Panis, S. (2020). How can we learn what attention is? Response gating via multiple direct
382 routes kept in check by inhibitory control processes. *Open Psychology*, 2(1), 238–279.
383 <https://doi.org/10.1515/psych-2020-0107>
- 384 Panis, S., Moran, R., Wolkersdorfer, M. P., & Schmidt, T. (2020). Studying the dynamics
385 of visual search behavior using RT hazard and micro-level speed–accuracy tradeoff
386 functions: A role for recurrent object recognition and cognitive control processes.
387 *Attention, Perception, & Psychophysics*, 82(2), 689–714.
388 <https://doi.org/10.3758/s13414-019-01897-z>
- 389 Panis, S., Schmidt, F., Wolkersdorfer, M. P., & Schmidt, T. (2020). Analyzing Response
390 Times and Other Types of Time-to-Event Data Using Event History Analysis: A Tool
391 for Mental Chronometry and Cognitive Psychophysiology. *I-Perception*, 11(6),
392 2041669520978673. <https://doi.org/10.1177/2041669520978673>

- 393 Panis, S., & Schmidt, T. (2016). What Is Shaping RT and Accuracy Distributions? Active
394 and Selective Response Inhibition Causes the Negative Compatibility Effect. *Journal of*
395 *Cognitive Neuroscience*, 28(11), 1651–1671. https://doi.org/10.1162/jocn_a_00998
- 396 Panis, S., & Schmidt, T. (2022). When does “inhibition of return” occur in spatial cueing
397 tasks? Temporally disentangling multiple cue-triggered effects using response history
398 and conditional accuracy analyses. *Open Psychology*, 4(1), 84–114.
399 <https://doi.org/10.1515/psych-2022-0005>
- 400 Panis, S., Torfs, K., Gillebert, C. R., Wagemans, J., & Humphreys, G. W. (2017).
401 Neuropsychological evidence for the temporal dynamics of category-specific naming.
402 *Visual Cognition*, 25(1-3), 79–99. <https://doi.org/10.1080/13506285.2017.1330790>
- 403 Panis, S., & Wagemans, J. (2009). Time-course contingencies in perceptual organization
404 and identification of fragmented object outlines. *Journal of Experimental Psychology:*
405 *Human Perception and Performance*, 35(3), 661–687.
406 <https://doi.org/10.1037/a0013547>
- 407 Pedersen, T. L. (2024). *Patchwork: The composer of plots*. Retrieved from
408 <https://CRAN.R-project.org/package=patchwork>
- 409 Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in s and s-PLUS*. New York:
410 Springer. <https://doi.org/10.1007/b98882>
- 411 R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna,
412 Austria: R Foundation for Statistical Computing. Retrieved from
413 <https://www.R-project.org/>
- 414 Schmidt, T., Panis, S., Wolkersdorfer, M. P., & Vorberg, D. (2022). Response inhibition in
415 the Negative Compatibility Effect in the absence of inhibitory stimulus features. *Open*
416 *Psychology*, 4(1), 219–230. <https://doi.org/10.1515/psych-2022-0012>
- 417 Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling*
418 *Change and Event Occurrence*. Oxford, New York: Oxford University Press.
- 419 Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design.

- 420 *Psychonomic Bulletin & Review*, 25(6), 2083–2101.
- 421 <https://doi.org/10.3758/s13423-018-1451-8>
- 422 Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical
423 distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, 108(3),
424 551–567. <https://doi.org/10.1037/0033-2909.108.3.551>
- 425 Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics.
426 *Acta Psychologica*, 41(1), 67–85. [https://doi.org/10.1016/0001-6918\(77\)90012-9](https://doi.org/10.1016/0001-6918(77)90012-9)
- 427 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New
428 York. Retrieved from <https://ggplot2.tidyverse.org>
- 429 Wickham, H. (2023a). *Forcats: Tools for working with categorical variables (factors)*.
430 Retrieved from <https://forcats.tidyverse.org/>
- 431 Wickham, H. (2023b). *Stringr: Simple, consistent wrappers for common string operations*.
432 Retrieved from <https://stringr.tidyverse.org>
- 433 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ...
434 Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43),
435 1686. <https://doi.org/10.21105/joss.01686>
- 436 Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A
437 grammar of data manipulation*. Retrieved from <https://dplyr.tidyverse.org>
- 438 Wickham, H., & Henry, L. (2023). *Purrr: Functional programming tools*. Retrieved from
439 <https://purrr.tidyverse.org/>
- 440 Wickham, H., Hester, J., & Bryan, J. (2024). *Readr: Read rectangular text data*. Retrieved
441 from <https://readr.tidyverse.org>
- 442 Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data*. Retrieved from
443 <https://tidyr.tidyverse.org>