

1 Event History Analysis for psychological time-to-event data: A tutorial in R with examples
2 in Bayesian and frequentist workflows

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). Social Brain Sciences lab, Department of Humanities, Social and Political
8 Sciences (D-GESS).

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

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

14

Abstract

15 Time-to-event data such as response times, saccade latencies, and fixation durations form a
16 cornerstone of experimental psychology, and have had a widespread impact on our
17 understanding of human cognition. However, the orthodox method for analysing such data
18 – comparing means between conditions – is known to conceal valuable information about
19 the timeline of psychological effects, such as their onset time and duration. The ability to
20 reveal finer-grained, “temporal states” of cognitive processes can have important
21 consequences for theory development by qualitatively changing the key inferences that are
22 drawn from psychological data. Moreover, well-established analytical approaches, such as
23 event history analysis, are able to evaluate the detailed shape of time-to-event distributions,
24 and thus characterise the timeline of psychological states. One barrier to wider use of event
25 history analysis, however, is that the analytical workflow is typically more time-consuming
26 and complex than orthodox approaches. To help achieve broader uptake, in this paper we
27 outline a set of tutorials that detail how to implement one distributional method known as
28 discrete-time event history analysis. We illustrate how to wrangle raw data files and
29 calculate descriptive statistics, as well as how to calculate inferential statistics via Bayesian
30 and frequentist multi-level regression modelling. We touch upon several key aspects of the
31 workflow, such as how to specify regression models, the implications for experimental
32 design, as well as how to manage inter-individual differences. We finish the article by
33 considering the benefits of the approach for understanding psychological states, as well as
34 the limitations and future directions of this work. Finally, the project is written in R and
35 freely available, which means the general approach can easily be adapted to other data
36 sets, and all of the tutorials are available as .html files to widen access beyond R-users.

37 *Keywords:* response times, event history analysis, Bayesian multi-level regression
38 models, experimental psychology, cognitive psychology

39 Word count: X

40 Event History Analysis for psychological time-to-event data: A tutorial in R with examples
41 in Bayesian and frequentist workflows

42 **1. Introduction**

43 **1.1 Motivation and background context: Comparing means versus
44 distributional shapes**

45 In experimental psychology, it is standard practice to analyse response times (RTs),
46 saccade latencies, and fixation durations by calculating average performance across a series
47 of trials. Such mean-average comparisons have been the workhorse of experimental
48 psychology over the last century, and have had a substantial impact on theory development
49 and our understanding of the structure of cognition and brain function. However,
50 differences in mean RT conceal when an experimental effect starts, how long it lasts, how it
51 evolves with increasing waiting time, and whether its onset is time-locked to other events
52 (Panis, 2020; Panis, Moran, Wolkersdorfer, & Schmidt, 2020; Panis & Schmidt, 2016, 2022;
53 Panis, Torfs, Gillebert, Wagemans, & Humphreys, 2017; Panis & Wagemans, 2009;
54 Wolkersdorfer, Panis, & Schmidt, 2020). Such information is useful not only for
55 interpretation of the effects, but also for cognitive psychophysiology and computational
56 model selection (Panis, Schmidt, Wolkersdorfer, & Schmidt, 2020).

57 As a simple illustration, Figure 1 shows the results of several simulated RT data sets,
58 which show how mean-average comparisons between two conditions can conceal the shape
59 of the underlying RT distributions. For instance, in examples 1-3, mean RT is always
60 comparable between two conditions, while the distributions differ (Figure 1, top row). In
61 contrast, in examples 4-6, mean RT is lower in condition 2 compared to condition 1, but
62 the RT distributions differ in each case (Figure 1, bottom row). Therefore, a comparison of
63 means would lead to a similar conclusion in examples 1-3, as well as examples 4-6, whereas
64 a comparison of the distributions would lead to a different conclusion in every case.

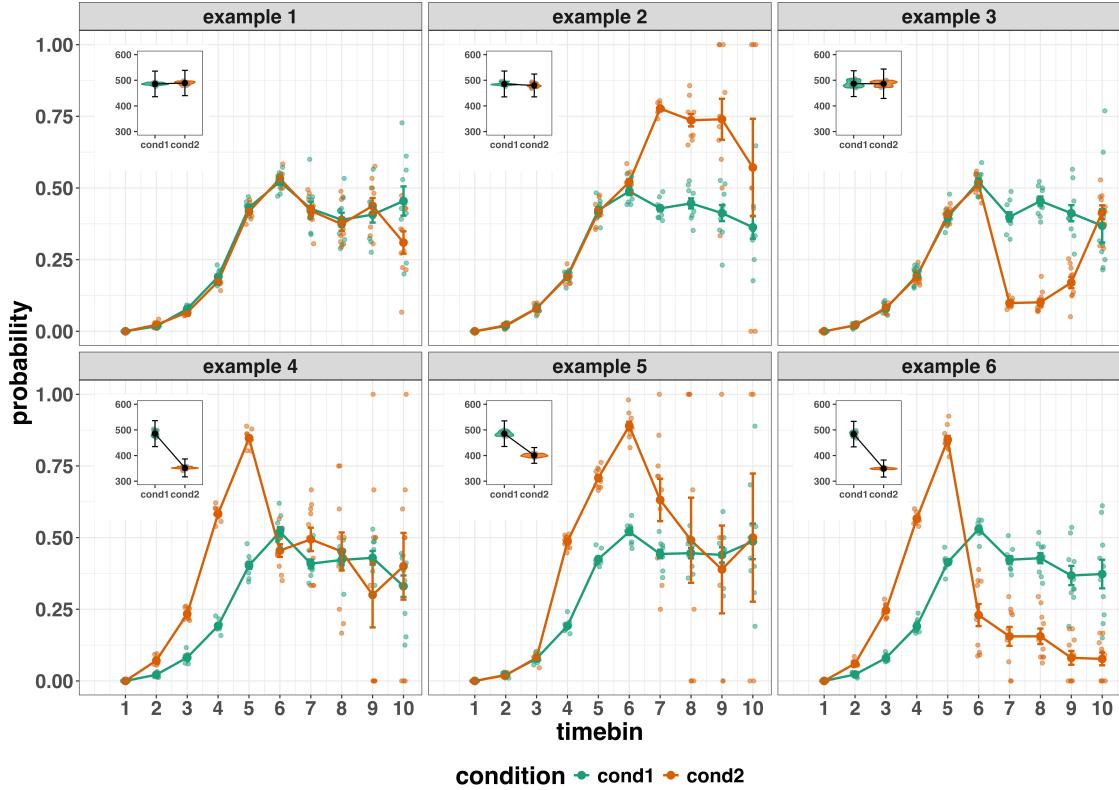


Figure 1. Means versus distributional shapes for six different simulated data set examples. The first second after stimulus onset is divided in ten bins of 100 ms. For our purposes here, it is enough to know that the distributions plotted represent the probability of an event occurring in that timebin, given that it has not yet occurred. Insets show mean reaction time per condition.

65 Why does this matter for research in psychology? Compared to the aggregation of
 66 data across trials, a distributional approach offers the possibility to reveal the time course
 67 of psychological states. As such, the approach permits different kinds of questions to be
 68 asked, different inferences to be made, and it holds the potential to discriminate between
 69 different theoretical accounts of psychological and/or brain-based processes. For example,
 70 the distributions in Example 4 show that the effect starts around 200 ms and is gone by
 71 600 ms. In contrast, in Example 5, the effect starts around 400 ms and is gone by 800 ms.
 72 And in the Example 6, the effect reverses around between 500 and 600 ms. What kind of

73 theory or set of theories could account for such effects? Are there new auxiliary
74 assumptions that theories need to adopt? And are there new experiments that need to be
75 run to test the novel predictions that follow from these analyses? As we show later using
76 concrete examples from past experimental data, for many psychological questions this
77 “temporal states” information can be theoretically meaningful by leading to more
78 fine-grained understanding of psychological processes as well as adding a relatively
79 under-used dimension – the passage of time – to our theory building toolkit.

80 From a historical perspective, it is worth noting that the development of analytical
81 tools that can estimate or predict whether and when events will occur is not a new
82 endeavour. Indeed, hundreds of years ago, analytical methods were developed to predict
83 time to death (e.g., William Matthew Makeham, 1860). The same logic has been applied
84 to psychological time-to-event data, as previously demonstrated (Panis et al., 2020). Here,
85 in the paper, we hope to show the value of event history analysis for knowledge and theory
86 building in cognitive psychology and related areas of research, such as cognitive
87 neuroscience, as well as provide practical tutorials that provide step-by-step code and
88 instructions in the hope that we can enable others to use event history analysis in a more
89 routine, efficient and effective manner.

90 1.2 Aims and structure of the paper

91 In this paper, we focus on a distributional method for time-to-event data known as
92 *discrete-time event history analysis*, a.k.a. hazard analysis, duration analysis, failure-time
93 analysis, survival analysis, and transition analysis. We first provide a brief overview of
94 event history analysis to orient the reader to the basic concepts that we will use
95 throughout the paper. However, this will remain relatively short, as this has been covered
96 in detail before (Allison, 1982, 2010; Singer & Willett, 2003), and our primary aim here is
97 to introduce a set of tutorials, which explain **how** to do such analyses, rather than repeat
98 in any detail **why** you should do them.

We then provide six different tutorials, each of which is written in the R programming language and publicly available on our Github and the Open Science Framework (OSF) pages, along with all of the other code and material associated with the project. The tutorials provide hands-on, concrete examples of key parts of the analytical process, so that others can apply the analyses to their own time-to-event data sets. Each tutorial is provided as an RMarkdown file, so that others can download and adapt the code to fit their own purposes. Additionally, each tutorial is made available as a .html file, so that it can be viewed by any web browser, and thus available to those that do not use R.

In Tutorial 1a, we illustrate how to process or “wrangle” a previously published RT + accuracy data set to calculate descriptive statistics when there is one independent variable. The descriptive statistics are plotted, and we comment on their interpretation. In Tutorial 1b we provide a generalisation of this approach to illustrate how one can calculate the descriptive statistics when using a more complex design, such as when there are two independent variables. In Tutorial 2a, we illustrate how one can fit Bayesian multi-level regression models to RT data using the R package brms. We discuss possible link functions, and plot the model-based effects of our predictors of interest. In Tutorial 2b we fit Bayesian multi-level regression models to *timed* accuracy data to perform a micro-level speed-accuracy tradeoff (SAT) analysis, which complements the event history analysis of RT data for choice RT data. In Tutorial 3a, we illustrate how to fit the same type of multilevel regression model for RT data in a frequentist framework using the R package lme4. We then briefly compare and contrast these inferential frameworks when applied to event history analysis. In Tutorial 3b, we illustrate how to perform the SAT analysis in a frequentist framework.

In summary, even though event history analysis is a widely used statistical tool and there already exist many excellent reviews (e.g., Blossfeld & Rohwer, 2002; Box-Steffensmeier, 2004; Hosmer, Lemeshow, & May, 2011; Teachman, 1983) and tutorials (e.g., Allison, 2010; Landes, Engelhardt, & Pelletier, 2020) on its general use-cases, we are

not aware of any tutorials that are aimed at psychological time-to-event data, and which provide worked examples of the key data processing and multi-level regression modelling steps. Therefore, our ultimate goal is twofold: first, we want to convince readers of the many benefits of using event history analysis when dealing with time-to-event data with a focus on psychological time-to-event data, and second, we want to provide a set of practical tutorials, which provide step-by-step instructions on how you actually perform a discrete-time event history analysis on time-to-event data such as RT data, as well as a complementary discrete-time SAT analysis on timed accuracy data.

2. A brief introduction to event history analysis

For a comprehensive background context to event history analysis, we recommend several excellent textbooks (Allison, 2010; Singer & Willett, 2003). Likewise, for general introduction to understanding regression equations, we recommend several introductory level textbooks (REFs). Our focus here is not on providing a detailed account of the underlying regression equations, since this topics has been comprehensively covered many times before. Instead, we want to provide an intuition to how event history analysis works in general as well as in the context of experimental psychology. As such, we only supply regression equations in the supplementary material (part D) and then refer to them in the text whenever relevant.

2.1 Basic features of event history analysis

To apply event history analysis (EHA), a.k.a. hazard analysis, one must be able to:

1. define an event of interest that represents a qualitative change that can be situated in time (e.g., a button press, a saccade onset, a fixation offset, etc.)
2. define time point zero (e.g., target stimulus onset, fixation onset)

- 149 3. measure the passage of time between time point zero and event occurrence in discrete
150 or continuous time units.

151 The definition of hazard and the type of models employed depend on whether one is
152 using continuous or discrete time units. Since our focus here is on hazard models that use
153 discrete time units, we describe that approach. After dividing time in discrete, contiguous
154 time bins indexed by t (e.g., $t = 1:10$ timebins), let RT be a discrete random variable
155 denoting the rank of the time bin in which a particular person's response occurs in a
156 particular experimental condition. For example, the first response might occur at 546 ms
157 and it would be in timebin 6 (any RTs from 501 ms to 600). Continuous RT data is treated
158 here as interval-censored data.

159 Discrete-time EHA focuses on the discrete-time hazard function of event occurrence
160 and the discrete-time survivor function (Figure 2). The equations that define both of these
161 functions are reported in the supplementary material (equations 1 and 2 in part A). The
162 discrete-time hazard function gives you for each bin the probability that the event occurs
163 (sometime) in bin t , given that the event does not occur in previous bins. In other words,
164 it reflects the instantaneous likelihood that the event occurs in the current bin, given that
165 it has not yet occurred in the past, i.e., in one of the prior bins. This conditionality in the
166 definition of hazard is what makes the hazard function so diagnostic for studying event
167 occurrence, as an event can physically not occur when it has already occurred before. In
168 contrast, the discrete-time survivor function cumulates the bin-by-bin risks of event
169 nonoccurrence to obtain the probability that the event occurs after bin t . In other words,
170 the survivor function gives you for each time bin the likelihood that the event occurs in the
171 future, i.e., in one of the subsequent timebins. As explained in part A of the supplementary
172 material, the survivor function provides a context for interpreting the hazard function.

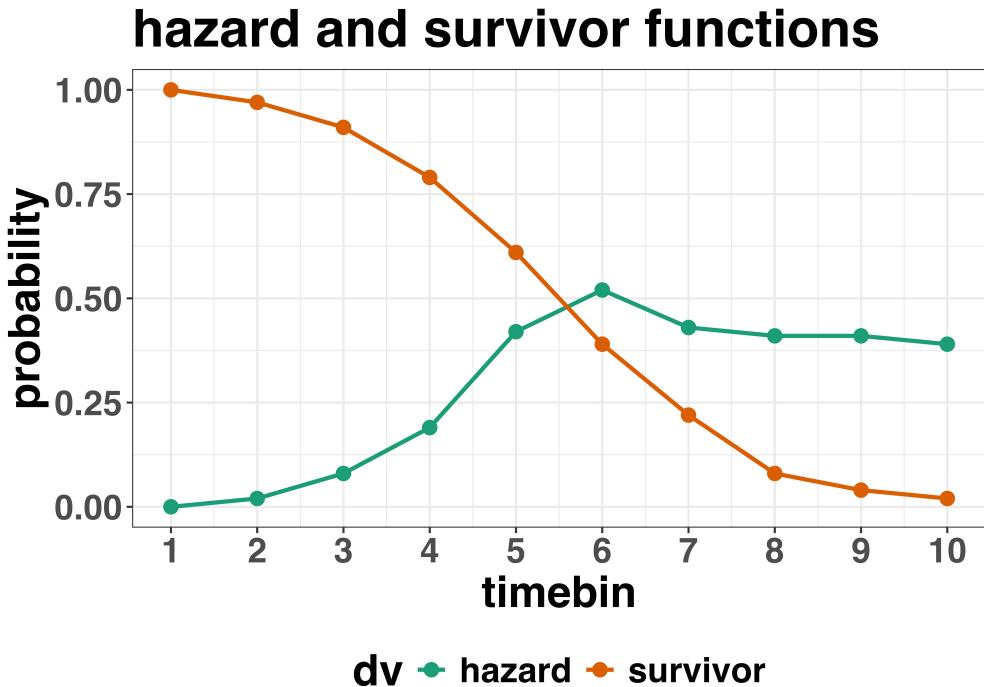


Figure 2. Discrete-time hazard and survivor functions.

¹⁷³ **2.2 Event history analysis in the context of experimental psychology**

¹⁷⁴ **2.2.1 A worked example.** In the context of experimental psychology, it is
¹⁷⁵ common for participants to be presented with either a 1-button detection task or a
¹⁷⁶ 2-button discrimination task, i.e., a task that has a right and a wrong answer. For
¹⁷⁷ example, a task may involve choosing between two response options with only one of them
¹⁷⁸ being correct. For such two-choice RT data, the discrete-time EHA can be extended with a
¹⁷⁹ discrete-time SAT analysis. Specifically, the hazard function of event occurrence can be
¹⁸⁰ extended with the discrete-time conditional accuracy function (see equation 5 in part A of
¹⁸¹ the supplementary material), which gives you the probability that a response is correct
¹⁸² given that it is emitted in time bin t (Allison, 2010; Kantowitz & Pachella, 2021;
¹⁸³ Wickelgren, 1977). We refer to this extended analysis for choice RT data as EHA/SAT. As
¹⁸⁴ explained in part A of the supplementary material, the probability mass function provides
¹⁸⁵ a context for interpreting the conditional accuracy function.

186 Integrating results between hazard and conditional accuracy functions for choice RT

187 data can be informative for understanding psychological processes. To illustrate, we

188 consider a hypothetical choice RT example that is inspired by real data (Panis & Schmidt,

189 2016), but simplified to make the main point clearer (Figure 3). In a standard response

190 priming paradigm, there is a prime stimulus (e.g., an arrow pointing left or right) followed

191 by a target stimulus (another arrow pointing left or right). The prime can then be

192 congruent or incongruent with the target. Figure 3 shows that the early upswing in hazard

193 is equal for both prime conditions, and that early responses are always correct in the

194 congruent condition and always incorrect in the incongruent condition. These results show

195 that for short waiting times (< bin 6), responses always follow the prime (and not the

196 target, as instructed). And then for longer waiting times, response hazard is lower in

197 incongruent compared to congruent trials, and all responses emitted in these later bins are

198 correct. This is interesting because mean-average RT would only represent the overall

199 ability of cognition to overcome interference, on average, across trials. And such a

200 conclusion is not supported when the effects are explored over a timeline. Instead, the

201 psychological conclusion is much more nuanced and suggests that multiple states start,

202 stop and possibly interact over a particular temporal window.

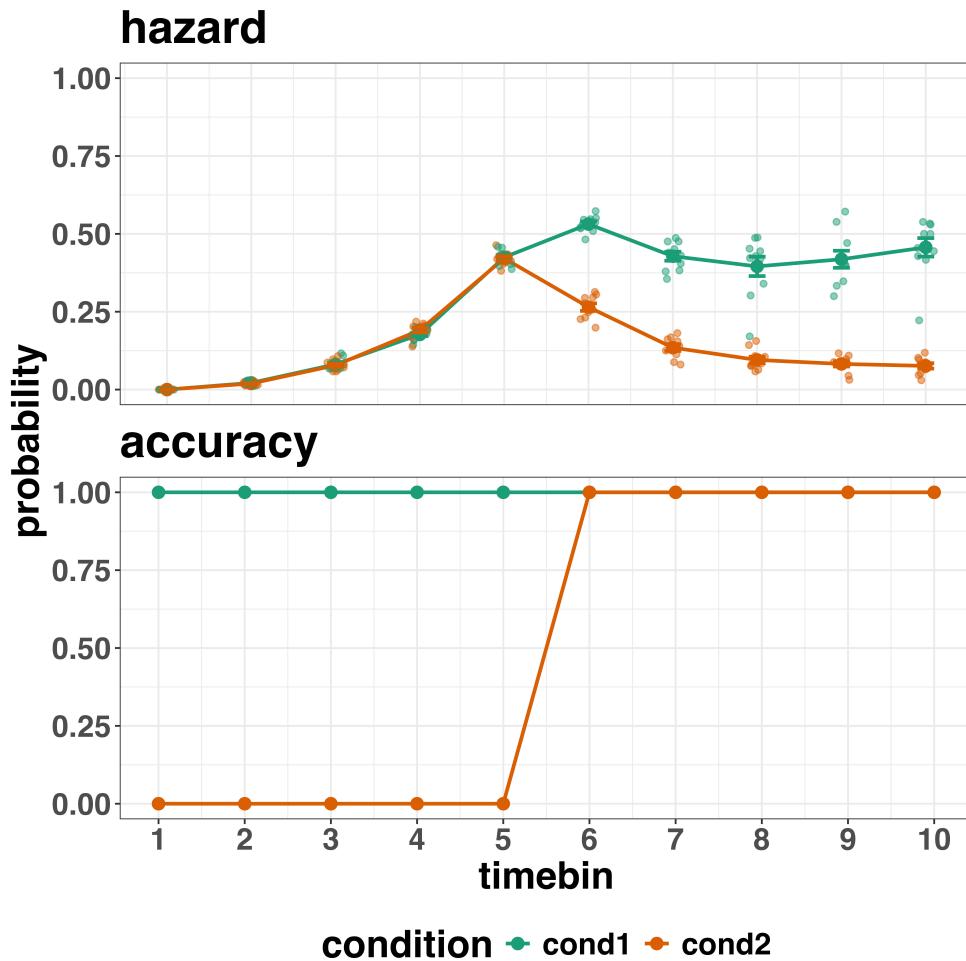


Figure 3. Discrete-time hazard and conditional accuracy functions.

Unlocking the temporal states of cognitive processes can be revealing in and of itself for theory development and the understanding of basic psychological processes. Possibly more importantly, however, is that it simultaneously opens the door to address many new and previously unanswered questions. Do all participants show similar temporal states or are there individual differences? Do such individual differences extend to those individuals that have been diagnosed with some form of psychopathology? How do temporal states relate to brain-based mechanisms that might be studied using other methods from cognitive neuroscience? And how much of theory in cognitive psychology would be in need of revision if mean-average comparisons were supplemented with a temporal states approach?

212 **2.2.2 Implications for designing experiments.** Performing event history

213 analyses in experimental psychology has implications for how experiments are designed.

214 Indeed, if trials are categorised as a function of when responses occur, then each timebin

215 will only include a subset of the total number of trials. For example, let's consider an

216 experiment where each participant performs 2 conditions and there are 100 trial repetitions

217 per condition. Those 100 trials must be distributed in some manner across the chosen

218 number of bins.

219 In such experimental designs, since the number of trials per condition are spread

220 across bins, it is important to have a relatively large number of trial repetitions per

221 participant and per condition. Accordingly, experimental designs using this approach

222 typically focus on factorial, within-subject designs, in which a large number of observations

223 are made on a relatively small number of participants (so-called small-*N* designs). This

224 approach emphasizes the precision and reproducibility of data patterns at the individual

225 participant level to increase the inferential validity of the design (Baker et al., 2021; Smith

226 & Little, 2018).

227 In contrast to the large-*N* design that typically average across many participants

228 without being able to scrutinize individual data patterns, small-*N* designs retain crucial

229 information about the data patterns of individual observers. This can be advantageous

230 whenever participants differ systematically in their strategies or in the time courses of their

231 effects, so that averaging them would lead to misleading data patterns. Note that because

232 statistical power derives both from the number of participants and from the number of

233 repeated measures per participant and condition, small-*N* designs can still achieve what

234 are generally considered acceptable levels of statistical power, if they have have a sufficient

235 amount of data overall (Baker et al., 2021; Smith & Little, 2018).

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

¹ We, furthermore, used the R-packages *bayesplot* (Version 1.11.1; Gabry, Simpson, Vehtari, Betancourt, &

²³⁷ content of the tutorials is mainly based on Allison (2010), Singer and Willett (2003),
²³⁸ McElreath (2018), Kurz (2023a), and Kurz (2023b).

²³⁹ **3. An overview of the general analytical workflow**

²⁴⁰ Although the focus is on EHA/SAT, we also want to briefly comment on broader
²⁴¹ aspects of our general analytical workflow, which relate more to data science and data
²⁴² analysis workflows.

²⁴³ **3.1 Data science workflow and descriptive statistics**

²⁴⁴ Descriptive, data science workflow. We perform data wrangling following tidyverse
²⁴⁵ principles and a functional programming approach (Wickham, Çetinkaya-Rundel, &
²⁴⁶ Grolemund, 2023). Functional programming basically means you don't write your own
²⁴⁷ loops but instead use functions that have been built and tested by others. [[more here, as
²⁴⁸ necessary]].

Gelman, 2019), *brms* (Version 2.21.0; Bürkner, 2017, 2018, 2021), *citr* (Version 0.3.2; Aust, 2019), *cmdstanr* (Version 0.8.1.9000; Gabry, Češnovar, Johnson, & Brander, 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), *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), *tidyR* (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).

249 3.2 Inferential statistical approach

250 Our lab adopts a estimation approach to multi-level regression (Kruschke & Liddell,
251 2018) ; Winter, 2019), which is heavily influenced by the Bayesian framework as suggested
252 by Richard McElreath (Kurz, 2023b; McElreath, 2018). We also use a “keep it maximal”
253 approach to specifying varying (or random) effects (Barr, Levy, Scheepers, & Tily, 2013).
254 This means that wherever possible we include varying intercepts and slopes per participant
255 To make inferences, we use two main approaches. We compare models of different
256 complexity, using information criteria (e.g., WAIC) and cross-validation (e.g., LOO), to
257 evaluate out-of-sample predictive accuracy (McElreath, 2018). We also take the most
258 complex model and evaluate key parameters of interest using point and interval estimates.

259 4. Tutorials

260 Tutorials 1a and 1b show how to calculate and plot the descriptive statistics of
261 EHA/SAT when there are one and two independent variables, respectively. Tutorials 2a
262 and 2b illustrate how to use Bayesian multilevel modeling to fit hazard and conditional
263 accuracy models, respectively. Tutorials 3a and 3b show how to implement, respectively,
264 multilevel models for hazard and conditional accuracy in the frequentist framework.
265 Additionally, to further simplify the process for other users, the tutorials rely on a set of
266 our own custom functions that make sub-processes easier to automate, such as data
267 wrangling and plotting functions (see part B in the supplemental material for a list of the
268 custom functions).

269 Our list of tutorials is as follows:

- 270 • 1a. Wrangle raw data and calculate descriptive stats for one independent variable.
- 271 • 1b. Wrangle raw data and calculate descriptive stats for two independent variables.
- 272 • 2a. Bayesian multilevel modeling for $h(t)$

- 273 • 2b. Bayesian multilevel modeling for $ca(t)$
- 274 • 3a. Frequentist multilevel modeling for $h(t)$
- 275 • 3b. Frequentist multilevel modeling for $ca(t)$

276 Planning (T4) - if we get a simulation and power analysis script working, which we

277 are happy with then we could include it here.

278 **4.1 Tutorial 1a: Calculating descriptive statistics using a life table**

279 **4.1.1 Data wrangling aims.** Our data wrangling procedures serve two related

280 purposes. First, we want to summarise and visualise descriptive statistics that relate to our
281 main research questions about the time course of psychological processes. Second, we want
282 to produce two different data sets that can each be submitted to different types of

283 inferential modelling approaches. The two types of data structure we label as ‘person-trial’
284 data and ‘person-trial-bin’ data. The ‘person-trial’ data (Table 1) will be familiar to most
285 researchers who record behavioural responses from participants, as it represents the
286 measured RT and accuracy per trial within an experiment. This data set is used when
287 fitting conditional accuracy models (Tutorials 2b and 3b).

Table 1
Data structure for ‘person-trial’ data

pid	trial	condition	rt	accuracy
1	1	congruent	373.49	1
1	2	incongruent	431.31	1
1	3	congruent	455.43	0
1	4	incongruent	622.41	1
1	5	incongruent	535.98	1
1	6	incongruent	540.08	1
1	7	congruent	511.07	1
1	8	incongruent	444.42	1
1	9	congruent	678.69	1
1	10	congruent	549.79	1

Note. The first 10 trials for participant 1 are shown. These data are simulated and for illustrative purposes only.

288 In contrast, the ‘person-trial-bin’ data (Table 2) has a different, more extended
 289 structure, which indicates in which bin a response occurred, if at all, in each trial.
 290 Therefore, the ‘person-trial-bin’ data set generates a 0 in each bin until an event occurs
 291 and then it generates a 1 to signal an event has occurred in that bin. This data set is used
 292 when fitting hazard models (Tutorials 2a and 3a). It is worth pointing out that there is no
 293 requirement for an event to occur at all (in any bin), as maybe there was no response on
 294 that trial or the event occurred after the time window of interest. Likewise, when the event
 295 occurs in bin 1 there would only be one row of data for that trial in the person-trial-bin
 296 data set.

Table 2
Data structure for ‘person-trial-bin’ data

pid	trial	condition	timebin	event
1	1	congruent	1	0
1	1	congruent	2	0
1	1	congruent	3	0
1	1	congruent	4	1
1	2	incongruent	1	0
1	2	incongruent	2	0
1	2	incongruent	3	0
1	2	incongruent	4	0
1	2	incongruent	5	1

Note. The first 2 trials for participant 1 from Table 1 are shown. The width of the time bins is 100 ms. These data are simulated and for illustrative purposes only.

297 **4.1.2 A real data wrangling example.** To illustrate how to quickly set up life
 298 tables for calculating the descriptive statistics (functions of discrete time), we use a
 299 published data set on masked response priming from Panis and Schmidt (2016). In their
 300 first experiment, Panis and Schmidt (2016) presented a double arrow for 94 ms that
 301 pointed left or right as the target stimulus with an onset at time point zero in each trial.
 302 Participants had to indicate the direction in which the double arrow pointed using their
 303 corresponding index finger, within 800 ms after target onset. Response time and accuracy
 304 were recorded on each trial. Prime type (blank, congruent, incongruent) and mask type
 305 were manipulated. Here we focus on the subset of trials in which no mask was presented.

306 The 13-ms prime stimulus was a double arrow with onset at -187 ms for the congruent
 307 (same direction as target) and incongruent (opposite direction as target) prime conditions.

308 There are several data wrangling steps to be taken. First, we need to load the data
 309 before we (a) supply required column names, and (b) specify the factor condition with the
 310 correct levels and labels.

311 The required column names are as follows:

- 312 • “pid”, indicating unique participant IDs;
- 313 • “trial”, indicating each unique trial per participant;
- 314 • “condition”, a factor indicating the levels of the independent variable (1, 2, ...) and
 315 the corresponding labels;
- 316 • “rt”, indicating the response times in ms;
- 317 • “acc”, indicating the accuracies (1/0).

318 In the code of Tutorial 1a, 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")))
```

319 Next, we can set up the life tables and plots of the discrete-time functions $h(t)$, $S(t)$,
 320 $ca(t)$, and $P(t)$ – see part A of the supplementary material for their definitions. To do so
 321 using a functional programming approach, one has to nest the data within participants
 322 using the group_nest() function, and supply a user-defined censoring time and bin width
 323 to our custom 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 data set

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

```

324 Note that the censoring time should be a multiple of the bin width (both in ms). The
 325 censoring time should be a time point after which no informative responses are expected
 326 anymore. In experiments that implement a response deadline in each trial the censoring
 327 time can equal that deadline time point. Trials with a RT larger than the censoring time,
 328 or trials in which no response is emitted during the data collection period, are treated as
 329 right-censored observations in EHA. In other words, these trials are not discarded, because
 330 they contain the information that the event did not occur before the censoring time.

331 Removing such trials before calculating the mean event time will result in underestimation
 332 of the true mean.

333 The person-trial-bin oriented data set is created by our custom function `ptb()`, and it
 334 has one row for each time bin (of each trial) that is at risk for event occurrence. The
 335 variable “event” in the person-trial-bin oriented data set indicates whether a response
 336 occurs (1) or not (0) for each bin.

337 The next step is to set up the life table using our custom function `setup_lt()`,
 338 calculate the conditional accuracies using our custom function `calc_ca()`, add the `ca(t)`
 339 estimates to the life table using our custom function `join_lt_ca()`, and then plot the
 340 descriptive statistics using our custom function `plot_eha()`. When creating the plots, some
 341 warning messages will likely be generated, like these:

- 342 • Removed 2 rows containing missing values or values outside the scale range
 343 `(geom_line())`.
- 344 • Removed 2 rows containing missing values or values outside the scale range
 345 `(geom_point())`.

- 346 • Removed 2 rows containing missing values or values outside the scale range
347 (`geom_segment()`).

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

352 In general, it is important to visually inspect the functions first for each participant,
353 in order to identify possible cheaters (e.g., a flat conditional accuracy function at .5
354 indicates (s)he was always guessing), outlying individuals, and/or different groups with
355 qualitatively different behavior.

356 Table 3 shows the life table for condition “blank” (no prime stimulus presented) for
357 participant 6. A life table includes for each time bin, the risk set (i.e., the number of trials
358 that are event-free at the start of the bin), the number of observed events, and the
359 estimates of $h(t)$, $S(t)$, $P(t)$, possibly $ca(t)$, and their estimated standard errors (se). At
360 time point zero, no events can occur and therefore $h(t)$ and $ca(t)$ are undefined.

361 Figure 4 displays the discrete-time hazard, survivor, conditional accuracy, and
362 probability mass functions for each prime condition for participant 6. By using
363 discrete-time hazard functions of event occurrence – in combination with conditional
364 accuracy functions for two-choice tasks – one can provide an unbiased, time-varying, and
365 probabilistic description of the latency and accuracy of responses based on all trials of any
366 data set.

367 For example, for participant 6, the estimated hazard values in bin (240,280] are 0.03,
368 0.17, and 0.20 for the blank, congruent, and incongruent prime conditions, respectively. In
369 other words, when the waiting time has increased until *240 ms* after target onset, then the
370 conditional probability of response occurrence in the next 40 ms is more than five times
371 larger for both prime-present conditions, compared to the blank prime condition.

372 Furthermore, the estimated conditional accuracy values in bin (240,280] are 0.29, 1,

373 and 0 for the blank, congruent, and incongruent prime conditions, respectively. In other

374 words, if a response is emitted in bin (240,280], then the probability that it is correct is

375 estimated to be 0.29, 1, and 0 for the blank, congruent, and incongruent prime conditions,

376 respectively.

Table 3

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	220	NA	NA	NA	1.00	0.00	NA	NA
40	220	0	0.00	0.00	1.00	0.00	NA	NA
80	220	0	0.00	0.00	1.00	0.00	NA	NA
120	220	0	0.00	0.00	1.00	0.00	NA	NA
160	220	0	0.00	0.00	1.00	0.00	NA	NA
200	220	0	0.00	0.00	1.00	0.00	NA	NA
240	220	0	0.00	0.00	1.00	0.00	NA	NA
280	220	7	0.03	0.01	0.97	0.01	0.29	0.17
320	213	13	0.06	0.02	0.91	0.02	0.77	0.12
360	200	26	0.13	0.02	0.79	0.03	0.92	0.05
400	174	40	0.23	0.03	0.61	0.03	1.00	0.00
440	134	48	0.36	0.04	0.39	0.03	0.98	0.02
480	86	37	0.43	0.05	0.22	0.03	1.00	0.00
520	49	32	0.65	0.07	0.08	0.02	1.00	0.00
560	17	9	0.53	0.12	0.04	0.01	1.00	0.00
600	8	4	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.

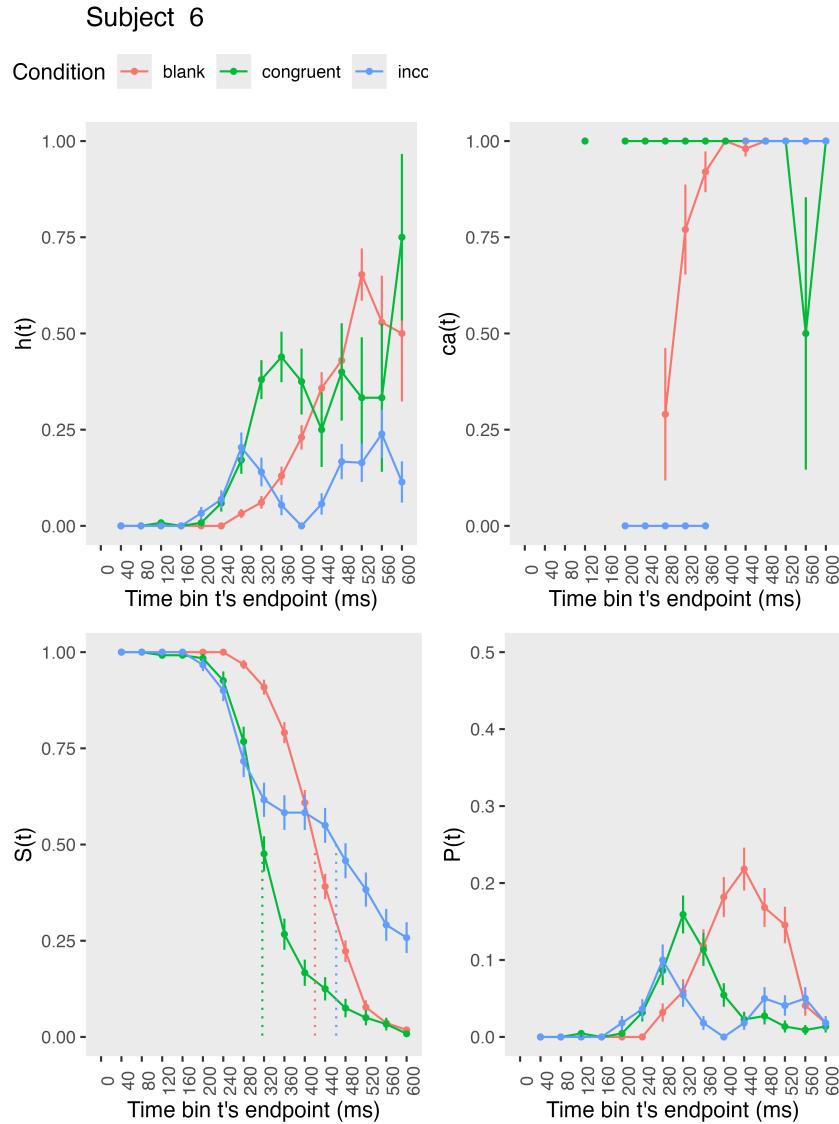


Figure 4. Estimated discrete-time hazard, survivor, probability mass, and conditional accuracy functions for participant 6, as a function of the passage of discrete waiting time.

377 However, when the waiting time has increased until 400 ms after target onset, then

378 the conditional probability of response occurrence in the next 40 ms is estimated to be

379 0.36, 0.25, and 0.06 for the blank, congruent, and incongruent prime conditions,

380 respectively. And when a response does occur in bin (400,440], then the probability that it

381 is correct is estimated to be 0.98, 1, and 1 for the blank, congruent, and incongruent prime

382 conditions, respectively.

383 When participants show qualitatively the same distributional patterns, one might
384 consider to aggregate their data and make one plot (see Tutorial_1a.Rmd).

385 These distributional results suggest that the participant 6 is initially responding to
386 the prime even though (s)he was instructed to only respond to the target, that response
387 competition emerges in the incongruent prime condition around 300 ms, and that only
388 slower responses are fully controlled by the target stimulus. Qualitatively similar results
389 were obtained for the other five participants.

390 In general, these results go against the (often implicit) assumption in research on
391 priming that all observed responses are primed responses to the target stimulus. Instead,
392 the distributional data show that early responses are triggered exclusively by the prime
393 stimulus, while only later responses reflect primed responses to the target stimulus.

394 At this point, we have calculated, summarised and plotted descriptive statistics for
395 the key variables in EHA/SAT. As we will show in later Tutorials, statistical models for
396 $h(t)$ and $ca(t)$ can be implemented as generalized linear mixed regression models predicting
397 event occurrence (1/0) and conditional accuracy (1/0) in each bin of a selected time
398 window for analysis. But first we consider calculating the descriptive statistics for two
399 independent variables.

400 4.2 Tutorial 1b: Generalising to a more complex design

401 So far in this paper, we have used a simple experimental design, which involved one
402 condition with three levels. But psychological experiments are often more complex, with
403 crossed factorial designs with more conditions and more than three levels. The purpose of
404 Tutorial 1b, therefore, is to provide a generalisation of the basic approach, which extends
405 to a more complicated design. We felt that this might be useful for researchers in
406 experimental psychology that typically use crossed factorial designs.

407 To this end, Tutorial 1b illustrates how to calculate and plot the descriptive statistics

408 for the full data set of Experiment 1 of Panis and Schmidt (2016), which includes two

409 independent variables: mask type and prime type. As we use the same functional

410 programming approach as in Tutorial 1a, we simply present the sample-based functions for

411 participant 6 in Figure 5.

412 In the no-mask condition (column 1 in Figure 5), we observe a positive compatibility

413 effect in the hazard and $ca(t)$ functions, as congruent primes temporarily generate higher

414 values for hazard and conditional accuracy compared to incongruent primes. However,

415 when a (relevant, irrelevant, or lines) mask is present (columns 2-4), there is a negative

416 compatibility effect in the hazard and conditional accuracy functions, as congruent primes

417 temporarily generate *lower* values for hazard and conditional accuracy compared to

418 incongruent primes.

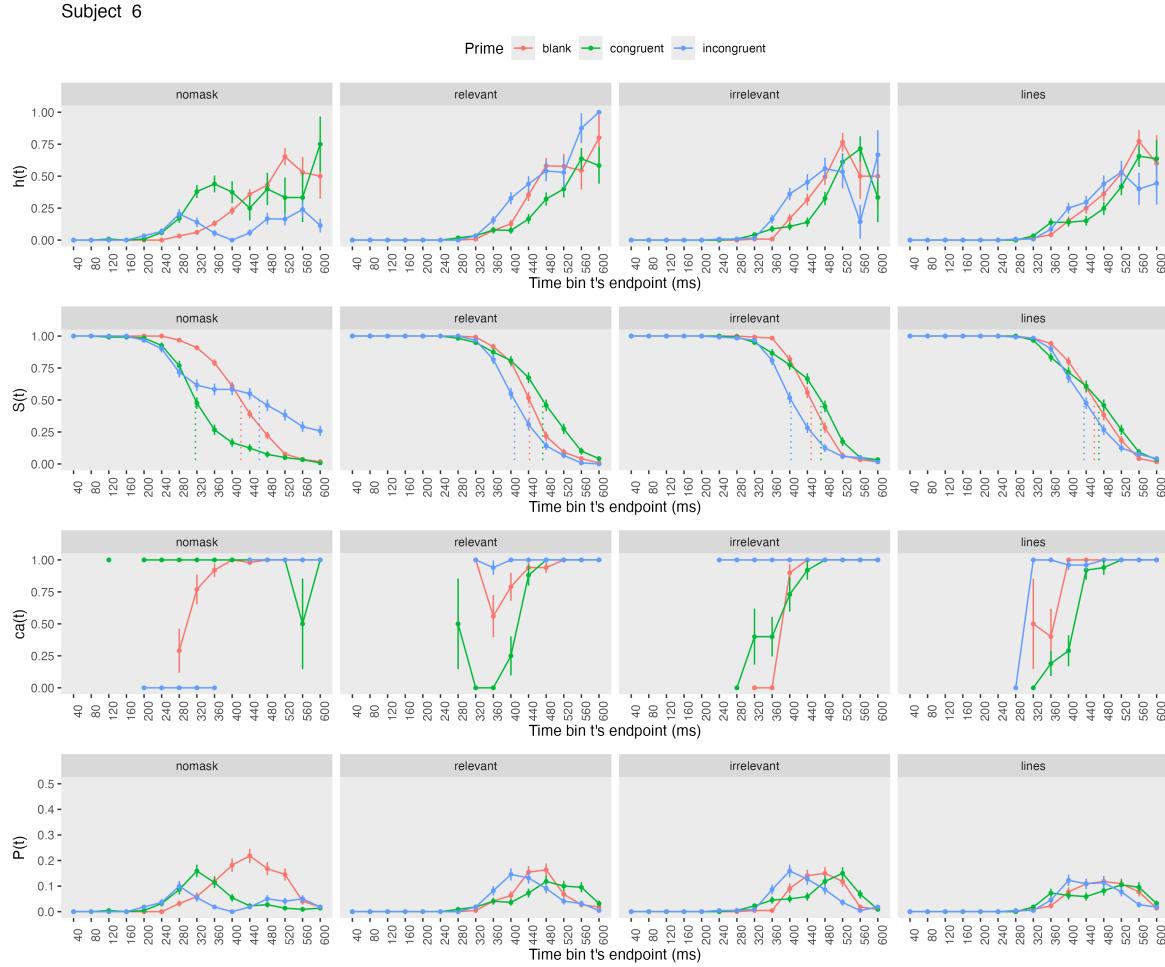


Figure 5. Sample-based discrete-time hazard, survivor, conditional accuracy, and probability mass functions for participant 6, as a function of prime type (blank, congruent, incongruent) and mask type (no mask, relevant, irrelevant, lines).

419 **4.3 Tutorial 2a: Fitting Bayesian hazard models to time-to-event data**

420 In this third tutorial, we illustrate how to fit Bayesian multi-level regression models
 421 to the RT data of the masked response priming data set used in Tutorial 1a. Fitting
 422 (Bayesian or non-Bayesian) regression models to time-to-event data is important when you
 423 want to study how the shape of the hazard function depends on various predictors (Singer
 424 & Willett, 2003).

425 **4.3.1 Hazard model considerations.** There are several analytic decisions one

426 has to make when fitting a hazard model. First, one has to select an analysis time window,
427 i.e., a contiguous set of bins for which there is enough data for each participant. Second,
428 given that the dependent variable (event occurrence) is binary, one has to select a link
429 function (see part C in the supplementary material). The cloglog link is preferred over the
430 logit link when events can occur in principle at any time point within a bin, which is the
431 case for RT data (Singer & Willett, 2003). Third, one has to choose a specification of the
432 effect of discrete TIME (i.e., the time bin index t) in a selected baseline condition. One can
433 choose a general specification (one intercept per bin) or a functional specification, such as a
434 polynomial one (compare model 1 with models 2, 3, and 4 below). Relevant example
435 regression formulas are provided in equations 6 and 7 in part D of the supplementary
436 material.

437 In the case of a large- N design without repeated measurements, the parameters of a

438 discrete-time hazard model can be estimated using standard logistic regression software
439 after expanding the typical person-trial data set into a person-trial-bin data set (Allison,
440 2010). When there is clustering in the data, as in the case of a small- N design with
441 repeated measurements, the parameters of a discrete-time hazard model can be estimated
442 using population-averaged methods (e.g., Generalized Estimating Equations), and Bayesian
443 or frequentist generalized linear mixed models (Allison, 2010).

444 In general, there are three assumptions one can make or relax when adding

445 experimental predictor variables and other covariates: The linearity assumption for
446 continuous predictors (the effect of a 1 unit change is the same anywhere on the scale), the
447 additivity assumption (predictors do not interact), and the proportionality assumption
448 (predictors do not interact with TIME).

449 In this tutorial we will fit four Bayesian multilevel models (i.e., generalized linear

450 mixed models) that differ in complexity to the person-trial-bin oriented data set that we

451 created in Tutorial 1a. We decided to select the analysis time window (200,600] and the
 452 cloglog link. The data is prepared as follows.

```
# load person-trial-bin oriented data set
ptb_data <- read_csv("../Tutorial_1_descriptive_stats/data/inputfile_hazard_modeling.csv")
# select analysis time window: (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 TIME (variable period) on bin 9, and variable trial on number 1000 and rescale; Add dummy variables for each bin.
ptb_data <- ptb_data %>%
  mutate(period_9 = period - 9,
         trial_c = (trial - 1000)/1000,
         d6 = if_else(period == 6, 1, 0),
         d7 = if_else(period == 7, 1, 0),
         d8 = if_else(period == 8, 1, 0),
         d9 = if_else(period == 9, 1, 0),
         d10 = if_else(period == 10, 1, 0),
         d11 = if_else(period == 11, 1, 0),
         d12 = if_else(period == 12, 1, 0),
         d13 = if_else(period == 13, 1, 0),
         d14 = if_else(period == 14, 1, 0),
         d15 = if_else(period == 15, 1, 0))
```

453 **4.3.2 Prior distributions.** To get the posterior distribution of each model

454 parameter given the data, we need to specify a prior distribution for each parameter. The
 455 middle column of Figure 12 in part E of the supplementary material shows seven examples
 456 of prior distributions on the logit and/or cloglog scales.

457 While a normal distribution with relatively large variance is often used as a weakly
 458 informative prior for continuous dependent variables, rows A and B in Figure 12 show that
 459 specifying such distributions on the logit and cloglog scales leads to rather informative
 460 distributions on the original probability scale, as most mass is pushed to probabilities of 0
 461 and 1. The other rows in Figure 12 show prior distributions on the logit and cloglog scale
 462 that we use instead.

463 **4.3.3 Model 1: A general specification of TIME, and main effects of**
 464 **congruency and trial number.** When you do not want to make assumptions about the
 465 shape of the hazard function in the selected baseline condition, or its shape is not smooth
 466 but irregular, then you can use a general specification of TIME, i.e., fit one intercept per
 467 time bin. In this first model, we use a general specification of TIME for the selected
 468 baseline condition (blank prime), and assume that the effects of prime-target congruency
 469 and trial number are proportional and additive, and that the effect of trial number is
 470 linear. Before we fit model 1, we remove unnecessary columns from the data, and specify
 471 our priors. In the code of Tutorial 2a, model M1 is specified as follows.

```
model_M1 <-  

  brm(data = M1_data,  

       family = binomial(link="cloglog"),  

       event | trials(1) ~ 0 + d6 + d7 + d8 + Intercept + d10 + d11 + d12 + d13 + d14 + d15 +  

              condition + trial_c +  

              (d6 + d7 + d8 + 1 + d10 + d11 + d12 + d13 + d14 + d15 +  

               condition + trial_c | pid),  

       prior = priors_M1,  

       chains = 4, cores = 4, iter = 3000, warmup = 1000,  

       control = list(adapt_delta = 0.999, step_size = 0.04, max_treedepth = 12),  

       seed = 12, init = "0",  

       file = "Tutorial_2_Bayesian/models/model_M1")
```

472 After selecting the binomial family and the cloglog link, the model formula is
 473 specified. The fixed effects include 9 dummy variables, the explicit Intercept variable
 474 (which represents bin 9 in this example), and the main effects of prime-target congruency
 475 (variable condition) and centered trial number (variable trial_c). Each of these effects is
 476 allowed to vary across individuals (variable pid). Estimating model M1 took about 70
 477 minutes on a MacBook Pro (Sonoma 14.6.1 OS, 18GB Memory, M3 Pro Chip).

478 **4.3.4 Model 2: A polynomial specification of TIME, and main effects of**
 479 **congruency and trial number.** When the shape of the hazard function is rather
 480 smooth, as it is for behavioral RT data, one can fit a more parsimonious model by using a

481 polynomial specification of TIME. For our second example model, we thus use a
 482 third-order polynomial specification of TIME for the selected baseline condition (blank
 483 prime), and again assume that the effects of prime-target congruency and centered trial
 484 number are proportional and additive, and that the effect of trial number is linear. The
 485 model formula for model M2 looks as follows.

```
event | trials(1) ~ 0 + Intercept + period_9 + I(period_9^2) + I(period_9^3) +
       condition + trial_c +
       (1 + period_9 + I(period_9^2) + I(period_9^3) +
       condition + trial_c | pid),
```

486 Because TIME is centered on bin 9, and trial number on trial 1000, the Intercept
 487 represents the cloglog-hazard in bin 9 for the blank prime condition in trial 1000.
 488 Estimating model M2 took about 2.5 hours.

489 **4.3.5 Model 3: A polynomial specification of TIME, and relaxing the
 490 proportionality assumption.** So far, we assumed that the effect of our predictors
 491 prime-target congruency and centered trial number are the same in each time bin. However,
 492 the descriptive plots (e.g., Figure 4) suggest that the effect of prime-target congruency
 493 varies across time bins. Previous research has shown that psychological effects typically
 494 change over time (Panis, 2020; Panis, Moran, et al., 2020; Panis & Schmidt, 2022; Panis et
 495 al., 2017; Panis & Wagemans, 2009). For the third model, we thus use a third-order
 496 polynomial specification of TIME for the baseline condition (blank prime), and relax the
 497 proportionality assumption for the predictor variables prime-target congruency (variable
 498 condition) and centered trial number (variable trial_c).

```
event | trials(1) ~ 0 + Intercept +
       condition*period_9 +
       condition*I(period_9^2) +
       condition*I(period_9^3) +
       trial_c*period_9 +
       trial_c*I(period_9^2) +
       trial_c*I(period_9^3) +
```

```
(1 + condition*period_9 +
condition*I(period_9^2) +
condition*I(period_9^3) +
trial_c*period_9 +
trial_c*I(period_9^2) +
trial_c*I(period_9^3) | pid),
```

499 Note that duplicate terms in the model formula are ignored. Estimating model M3

500 took about 4.5 hours.

501 4.3.6 Model 4: A polynomial specification of TIME, and relaxing all three

502 assumptions. Based on previous work (e.g., Panis, 2020) we expect nonlinear effects of

503 trial number on the hazard of response occurrence. We thus relax all three assumptions in

504 model 4. We add a squared term for the continuous predictor centered trial number –

505 $I(trial_c^2)$ – and include interaction terms. For example, how the effect of congruent

506 primes changes across time bins within a trial might change across the trials within an

507 experiment.

```
event | trials(1) ~ 0 + Intercept +
      condition*period_9*trial_c +
      condition*period_9*I(trial_c^2) +
      condition*I(period_9^2)*trial_c +
      condition*I(period_9^2)*I(trial_c^2) +
      condition*I(period_9^3) +
      trial_c*I(period_9^3) +
      (1 + condition*period_9*trial_c +
      condition*period_9*I(trial_c^2) +
      condition*I(period_9^2)*trial_c +
      condition*I(period_9^2)*I(trial_c^2) +
      condition*I(period_9^3) +
      trial_c*I(period_9^3) | pid)
```

508 Again, duplicate terms in the model formula are ignored. Estimating model M4 took

509 about 8 hours.

510 **4.3.7 Compare the models.** We can compare the four models using the Widely

511 Applicable Information Criterion (WAIC) and Leave-One-Out (LOO) cross-validation, and
 512 look at model weights for both criteria (Kurz, 2023a; McElreath, 2018).

```
model_weights(model_M1, model_M2, model_M3, model_M4, weights = "loo") %>% round(digits = 3)
```

```
513 ## model_M1 model_M2 model_M3 model_M4
514 ##          0      0      0      1
```

```
model_weights(model_M1, model_M2, model_M3, model_M4, weights = "waic") %>% round(digits = 3)
```

```
515 ## model_M1 model_M2 model_M3 model_M4
516 ##          0      0      0      1
```

517 Clearly, both the loo and waic weighting schemes assign a weight of 1 to model M4,
 518 and a weight of 0 to the other three simpler models.

519 **4.3.8 Evaluate parameter estimates.** To make inferences from the parameter
 520 estimates in model M4, we summarize the draws from the posterior distributions of the
 521 effects of congruent and incongruent primes relative to the blank prime condition, in each
 522 time bin for trial numbers 500, 1000, and 1500, in terms of point and interval estimates.

523 Figure 6 shows one point (mean) and three highest posterior density interval
 524 (50/80/95%) estimates for the effects of congruent and incongruent primes relative to
 525 neutral primes, for each time bin in trial numbers 500, 1000, and 1500.

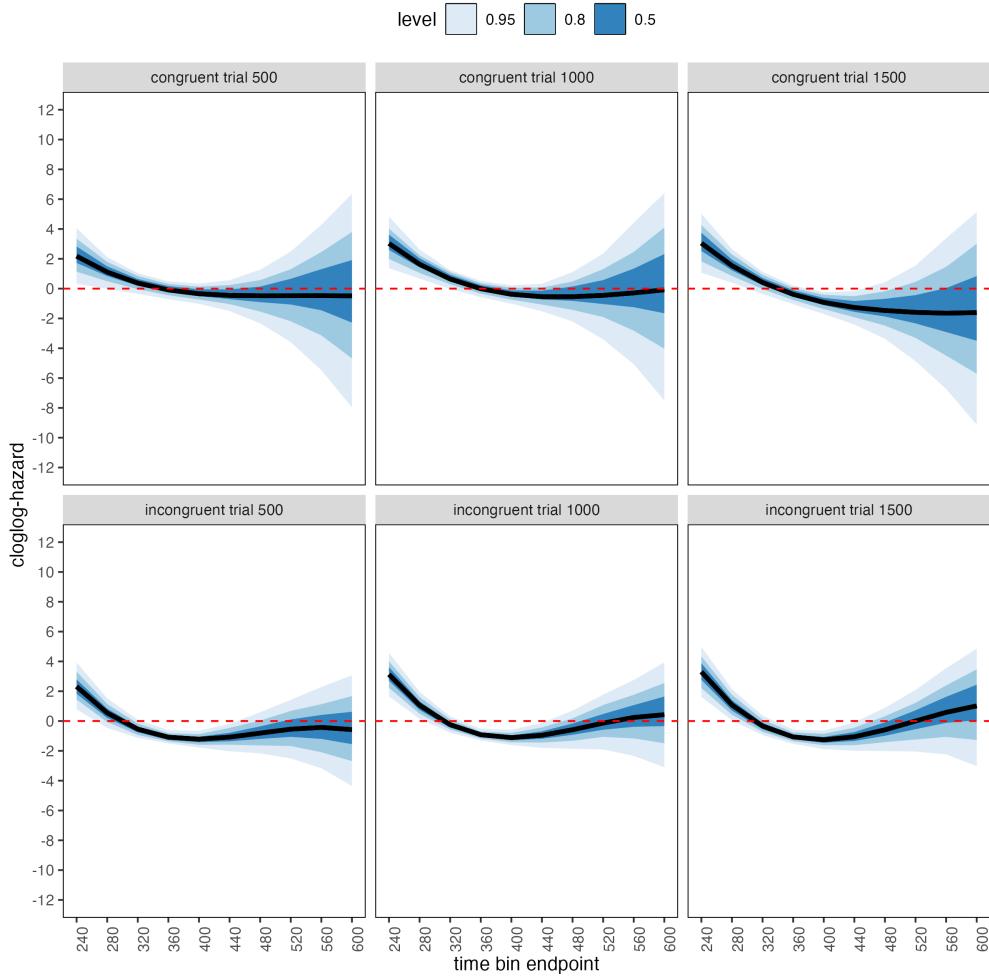


Figure 6. Means and 50/80/95% highest posterior density intervals of the draws from the posterior distributions representing the effect of congruent and incongruent primes relative to neutral primes in each bin for trial numbers 500, 1000, and 1500.

526 Table 4 shows the summaries of the draws from the posterior distributions of the
 527 effects of congruent and incongruent primes relative to the blank prime condition in trials
 528 500, 1000, and 1500, in terms of a point estimate (the mean) and the upper and lower
 529 bounds of the 95% highest posterior density interval. Also, by exponentiating the mean we
 530 obtain an effect size in terms of a hazard ratio.

Table 4

*Point and 95% highest posterior density interval estimates,
and hazard ratios.*

bin_endpoint	condition	mean	.lower	.upper	.width	hazard ratio
240	c500	2.18	0.33	4.05	0.95	8.82080
280	c500	1.11	-0.02	2.11	0.95	3.03199
320	c500	0.37	-0.34	1.04	0.95	1.45383
360	c500	-0.09	-0.70	0.48	0.95	0.91397
400	c500	-0.35	-1.02	0.34	0.95	0.70802
440	c500	-0.45	-1.50	0.56	0.95	0.63522
480	c500	-0.48	-2.32	1.27	0.95	0.62035
520	c500	-0.48	-3.57	2.52	0.95	0.61980
560	c500	-0.52	-5.69	4.27	0.95	0.59543
600	c500	-0.66	-8.56	6.99	0.95	0.51694
240	c1000	3.03	1.37	4.82	0.95	20.63183
280	c1000	1.63	0.68	2.63	0.95	5.12611
320	c1000	0.64	-0.02	1.24	0.95	1.90342
360	c1000	-0.01	-0.57	0.52	0.95	0.99277
400	c1000	-0.38	-1.01	0.22	0.95	0.68359
440	c1000	-0.54	-1.52	0.32	0.95	0.58403
480	c1000	-0.54	-2.20	1.11	0.95	0.58190
520	c1000	-0.45	-3.40	2.35	0.95	0.63546
560	c1000	-0.34	-5.78	3.90	0.95	0.71487
600	c1000	-0.25	-8.34	6.73	0.95	0.77863
240	c1500	3.05	1.07	5.02	0.95	21.02227
280	c1500	1.54	0.40	2.65	0.95	4.65584
320	c1500	0.42	-0.36	1.13	0.95	1.51502
360	c1500	-0.38	-1.05	0.21	0.95	0.68077
400	c1500	-0.92	-1.70	-0.24	0.95	0.39703
440	c1500	-1.26	-2.41	-0.18	0.95	0.28245
480	c1500	-1.47	-3.36	0.43	0.95	0.23037
520	c1500	-1.60	-4.86	1.58	0.95	0.20247
560	c1500	-1.71	-7.01	3.37	0.95	0.18021
600	c1500	-1.88	-10.07	5.98	0.95	0.15267
240	i500	2.31	0.79	3.93	0.95	10.10461
280	i500	0.55	-0.46	1.52	0.95	1.72468
320	i500	-0.54	-1.13	0.08	0.95	0.58233
360	i500	-1.08	-1.50	-0.61	0.95	0.33902
400	i500	-1.22	-1.78	-0.65	0.95	0.29661

Table 4 continued

bin_endpoint	condition	mean	.lower	.upper	.width	hazard ratio
440	i500	-1.08	-2.03	-0.19	0.95	0.33991
480	i500	-0.81	-2.16	0.59	0.95	0.44474
520	i500	-0.55	-2.50	1.42	0.95	0.57904
560	i500	-0.42	-3.16	2.28	0.95	0.65388
600	i500	-0.58	-4.35	3.10	0.95	0.55820
240	i1000	3.12	1.66	4.58	0.95	22.68463
280	i1000	1.06	0.15	1.95	0.95	2.88377
320	i1000	-0.24	-0.78	0.31	0.95	0.78490
360	i1000	-0.92	-1.30	-0.52	0.95	0.39866
400	i1000	-1.11	-1.61	-0.59	0.95	0.32935
440	i1000	-0.95	-1.80	-0.12	0.95	0.38574
480	i1000	-0.58	-1.86	0.70	0.95	0.55825
520	i1000	-0.14	-1.90	1.77	0.95	0.87013
560	i1000	0.24	-2.33	2.75	0.95	1.27313
600	i1000	0.42	-3.17	3.85	0.95	1.52411
240	i1500	3.30	1.63	4.98	0.95	27.07329
280	i1500	1.08	0.05	2.14	0.95	2.93821
320	i1500	-0.33	-0.94	0.36	0.95	0.71847
360	i1500	-1.06	-1.52	-0.57	0.95	0.34502
400	i1500	-1.26	-1.88	-0.65	0.95	0.28360
440	i1500	-1.06	-1.99	-0.09	0.95	0.34778
480	i1500	-0.59	-2.01	0.88	0.95	0.55459
520	i1500	0.00	-2.05	2.09	0.95	1.00234
560	i1500	0.58	-2.23	3.54	0.95	1.78962
600	i1500	1.01	-3.02	4.86	0.95	2.75123

Note. c500 = effect of congruent primes in trial 500; c1000 = effect of congruent primes in trial 1000; c1500 = effect of congruent primes in trial 1500; i500 = effect of incongruent primes in trial 500; i1000 = effect of incongruent primes in trial 1000; i1500 = effect of incongruent primes in trial 1500.

532 Based on Figure 6 and Table 4, we see that at the beginning of the experiment (trial
533 500), congruent and incongruent primes have a positive effect in time bin (200,240] on
534 cloglog-hazard, relative to the cloglog-hazard estimate in the baseline condition (no prime;
535 red striped lines in Figure 6). For example, the hazard ratio shows that the hazard of
536 response occurrence for congruent primes is estimated to be 8.82 times higher than that for
537 no-prime trials in bin (200,240] of trial 500. Incongruent primes also have a negative effect
538 on cloglog-hazard in bins (320,360], (360,400], and (400,440]. For example, in bin (320,360],
539 the hazard ratio shows that the hazard of response occurrence for incongruent prime is
540 estimated to be .34 times smaller than that for no-prime trials. While the early positive
541 effects reflect responses to the prime stimulus, the later negative effect for incongruent
542 primes likely reflects response competition between the prime-triggered response (e.g., left)
543 and the target-triggered response (e.g., right)

544 In the middle of the experiment (trial 1000), both congruent and incongruent primes
545 have positive effects in bins (200,240] and (240,280], while incongruent primes again have
546 negative effects in bins (320,360], (360,400], and (400,440]. Probably due to practicing
547 stimulus-response associations, the primes generate a higher hazard of response occurrence
548 for 80 ms early in a trial (compared to 40 ms at the beginning of the experiment)
549 compared to the blank prime condition.

550 Towards the end of the experiment (trial 1500), both congruent and incongruent
551 primes have positive and negative effects. Positive effects are present in bins (200,240] and
552 (240,280]. Incongruent primes again have negative effects in bins (320,360], (360,400], and
553 (400,440], and congruent primes now also have negative effects in bins (360,400] and
554 (400,440].

555 These results show that the effect of prime-target congruency changes not only on the
556 across-bin/within-trial time scale (variable period_9), but also on the
557 across-trial/within-experiment time scale (variable trial_c). The fact that congruent

558 primes generate negative effects for 80 ms (compared to no-prime trials) towards the end of
559 the experiment, while incongruent primes generate negative effects for 120 ms throughout
560 the experiment, strongly suggests the involvement of separate cognitive processes.

561 Panis and Schmidt (2016) distinguished between automatic response competition
562 (bottom-up lateral inhibition between response channels), active and global inhibition
563 (top-down nonselective response inhibition), and active and selective inhibition (top-down
564 selective response inhibition). While automatic response competition can be expected to be
565 present in the incongruent trials throughout the experiment, active and global response
566 inhibition effects might be present in both congruent and incongruent (unmasked) prime
567 trials. In other words, people learn that the prime-triggered response is premature and that
568 they have to temporarily slow down (increase the global response threshold) in order to
569 allow gating of the correct response to the target stimulus. Thus, it seems that this global
570 inhibitory effect becomes visible in the congruent (compared to no-prime) trials towards
571 the end of the experiment, while it might be masked by the automatic inhibitory effect of
572 response competition in the incongruent trials. Interestingly, while Panis and Schmidt
573 (2016) did not test interactions between prime-target congruency and trial number, they
574 concluded that active (i.e., top-down) response inhibition starts around 360 ms after the
575 onset of the second stimulus (the target stimulus in no-mask trials), which nicely coincides
576 with the onset of the negative effect of congruent primes observed here in trial 1500.

577 To conclude this Tutorial 2a, Figure 7 shows the model-based hazard functions for
578 each prime type for participant 6, in trial 500, 1000, and 1500.

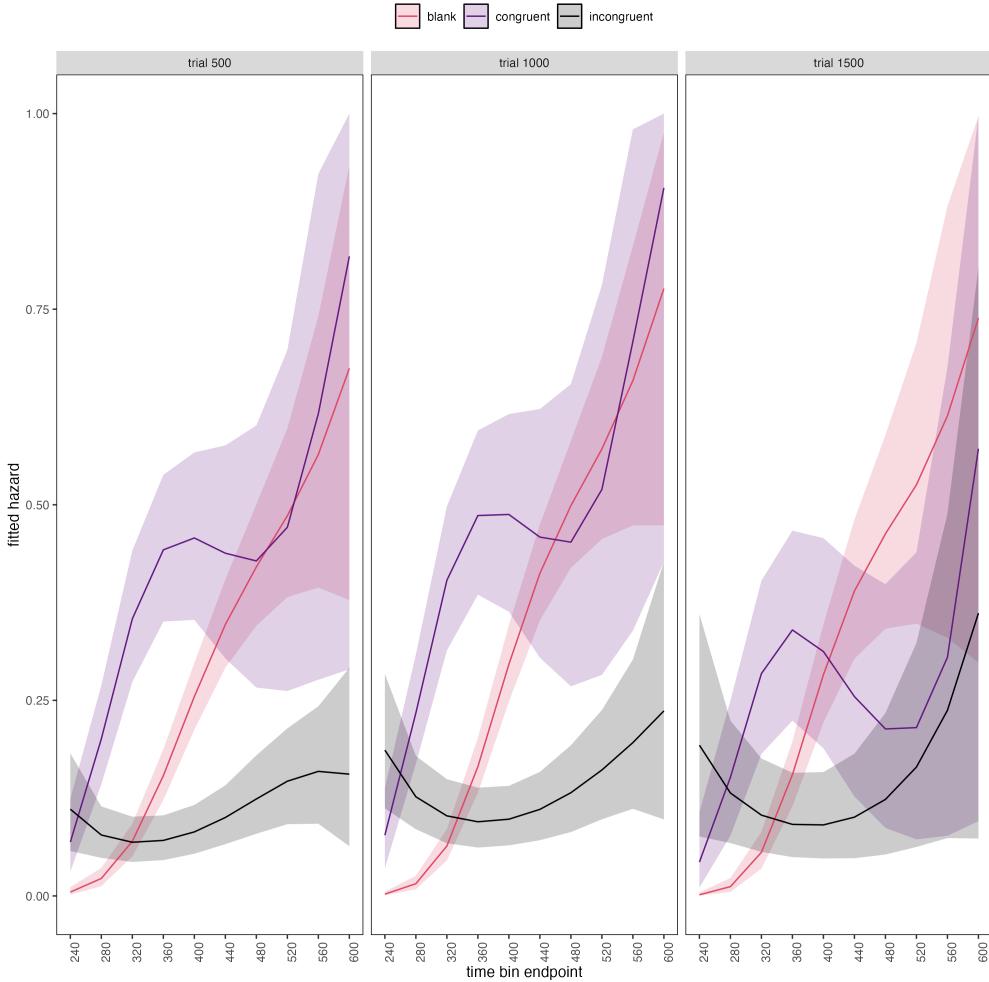


Figure 7. Model-based hazard functions for participant 6 in trial 500, 1000, and 1500.

579 **4.4 Tutorial 2b: Fitting Bayesian conditional accuracy models**

580 In this fourth tutorial, we illustrate how to fit a Bayesian multi-level regression model
 581 to the timed accuracy data from the masked response priming data set used in Tutorial 1a.
 582 The general process is similar to Tutorial 2a, except that (a) we use the person-trial data
 583 set, (b) we use the logit link function, and (c) we change the priors. For illustration
 584 purposes, we only fitted the effects of model M4 (see Tutorial 2a) in the conditional
 585 accuracy model called M4_ca.

586 To make inferences from the parameter estimates in model M4_ca, we summarize the

587 draws from the posterior distributions of the effects of congruent and incongruent primes
 588 on logit-ca relative to the blank prime condition, in each time bin for trial numbers 500,
 589 1000, and 1500, in terms of point and interval estimates.

590 Figure 8 shows one point (mean) and three highest posterior density interval
 591 (50/80/95%) estimates for the effects of congruent and incongruent primes relative to
 592 neutral primes on logit-ca, for each time bin in trial numbers 500, 1000, and 1500.

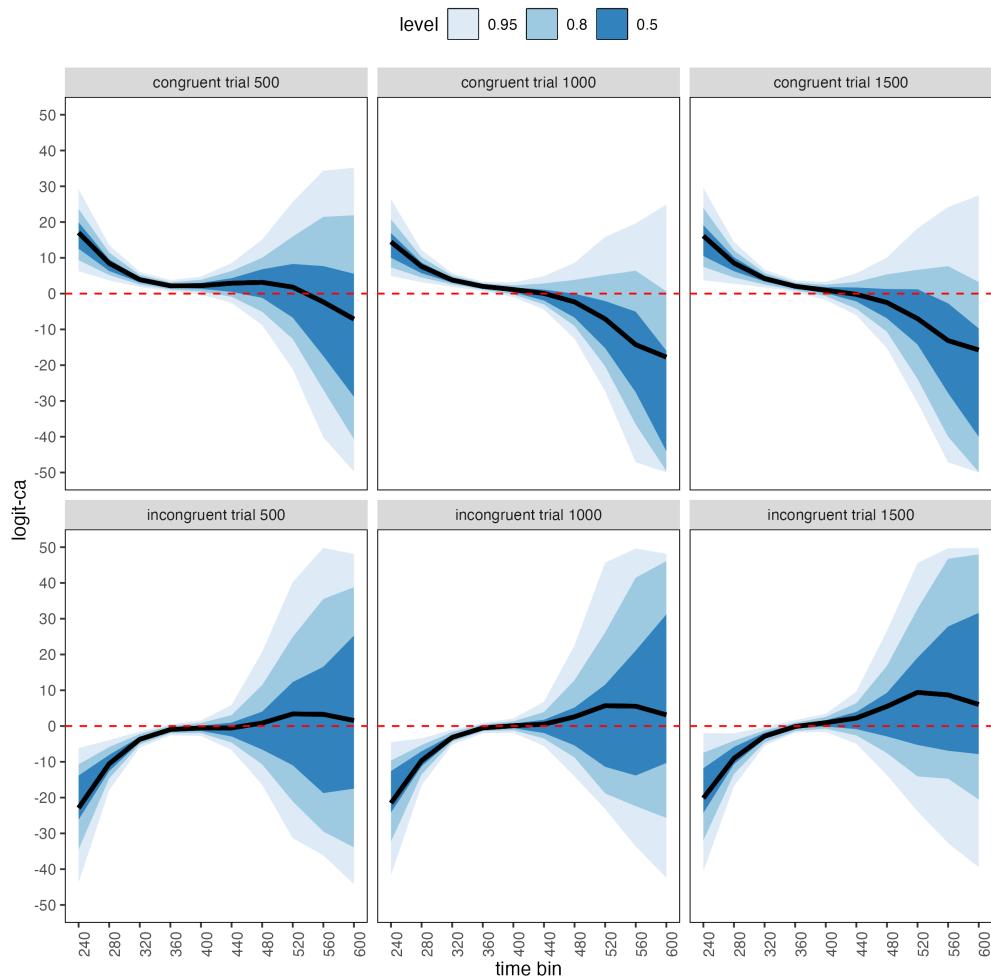


Figure 8. Means and 50/80/95% highest posterior density intervals of the draws from the posterior distributions representing the effect of congruent and incongruent primes relative to neutral primes in each bin for trial numbers 500, 1000, and 1500.

593 Table 5 shows the summaries of the draws from the posterior distributions of the

594 effects of congruent and incongruent primes relative to the blank prime condition in trials
 595 500, 1000, and 1500, in terms of a point estimate (the mean) and the upper and lower
 596 bounds of the 95% highest posterior density interval. Also, by exponentiating the mean we
 597 obtain an effect size in terms of an odds ratio.

Table 5

Point and 95% highest posterior density interval estimates, and odds ratios.

bin_endpoint	condition	mean	.lower	.upper	.width	odds ratio
240	c500	17.02	6.26	29.22	0.95	24618458.6117376089096
280	c500	8.49	3.71	13.54	0.95	4846.0911175084821
320	c500	3.91	1.88	6.05	0.95	49.7913398591541
360	c500	2.19	0.69	3.75	0.95	8.8918515266144
400	c500	2.22	-0.25	4.75	0.95	9.1907589464532
440	c500	2.91	-2.56	8.54	0.95	18.3094780953354
480	c500	3.15	-8.77	15.10	0.95	23.4104953301397
520	c500	1.86	-20.13	26.73	0.95	6.3973471906322
560	c500	-2.08	-39.94	42.41	0.95	0.1244195286526
600	c500	-9.77	-73.17	61.54	0.95	0.0000573483764
240	c1000	14.46	4.94	26.35	0.95	1899836.0176862408407
280	c1000	7.58	3.21	12.18	0.95	1961.8278183493471
320	c1000	3.80	1.90	5.71	0.95	44.8743225515544
360	c1000	2.02	0.72	3.35	0.95	7.5713234449885
400	c1000	1.14	-0.99	3.11	0.95	3.1378140505531
440	c1000	0.06	-4.41	4.87	0.95	1.0636745027736
480	c1000	-2.32	-12.62	8.61	0.95	0.0982112115060
520	c1000	-7.10	-27.24	15.97	0.95	0.0008224865296
560	c1000	-15.39	-54.71	23.54	0.95	0.0000002080447
600	c1000	-28.27	-92.96	35.54	0.95	0.0000000000005
240	c1500	16.12	3.74	29.48	0.95	10001085.3946626689285
280	c1500	8.54	2.78	14.43	0.95	5124.4351045810508
320	c1500	4.22	1.75	6.70	0.95	68.1208056719069
360	c1500	2.06	0.48	3.71	0.95	7.8233561607888
400	c1500	0.95	-1.75	3.26	0.95	2.5848218057173
440	c1500	-0.20	-6.03	5.65	0.95	0.8181575226781
480	c1500	-2.49	-15.23	10.07	0.95	0.0826145785621
520	c1500	-7.03	-30.41	18.55	0.95	0.0008862056864
560	c1500	-14.91	-58.81	27.21	0.95	0.0000003362909
600	c1500	-27.22	-95.59	43.50	0.95	0.00000000000015

Table 5 continued

bin_endpoint	condition	mean	.lower	.upper	.width	odds ratio
240	i500	-23.34	-44.42	-4.87	0.95	0.0000000000730
280	i500	-10.55	-17.93	-3.94	0.95	0.0000261643171
320	i500	-3.71	-6.06	-1.51	0.95	0.0246001196011
360	i500	-0.97	-2.57	0.56	0.95	0.3780876098423
400	i500	-0.52	-2.75	1.55	0.95	0.5916606671310
440	i500	-0.53	-6.67	5.86	0.95	0.5871858652992
480	i500	0.83	-16.41	20.71	0.95	2.3019651321687
520	i500	5.40	-32.44	52.48	0.95	222.0443659277399
560	i500	15.00	-58.75	104.35	0.95	3282435.9279020344839
600	i500	31.47	-90.20	190.08	0.95	46319712352328.7578125000000
240	i1000	-21.85	-43.05	-4.10	0.95	0.0000000003243
280	i1000	-9.67	-16.56	-3.46	0.95	0.0000632158160
320	i1000	-3.17	-5.23	-0.99	0.95	0.0419655563481
360	i1000	-0.53	-2.03	0.89	0.95	0.5909004105316
400	i1000	0.09	-1.88	2.11	0.95	1.0992267336787
440	i1000	0.52	-5.54	6.73	0.95	1.6827111411806
480	i1000	2.58	-14.16	22.53	0.95	13.2031868705690
520	i1000	8.10	-28.51	55.88	0.95	3307.4439707159477
560	i1000	18.92	-51.75	111.96	0.95	164758701.8390493392944
600	i1000	36.86	-89.39	191.12	0.95	10165856639901592.0000000000000
240	i1500	-20.51	-42.95	-2.49	0.95	0.0000000012362
280	i1500	-9.04	-16.80	-2.03	0.95	0.0001189822174
320	i1500	-2.86	-5.47	-0.25	0.95	0.0575421441866
360	i1500	-0.14	-1.81	1.67	0.95	0.8709638702927
400	i1500	0.94	-1.61	3.40	0.95	2.5699339941536
440	i1500	2.22	-4.97	9.63	0.95	9.2076572160001
480	i1500	5.52	-13.87	26.57	0.95	249.5050127299390
520	i1500	12.67	-31.46	58.41	0.95	318500.3986836019321
560	i1500	25.50	-53.08	115.21	0.95	119299568240.9411773681641
600	i1500	45.85	-86.60	200.06	0.95	81670189671651033088.0000000000000

Note. c500 = effect of congruent primes in trial 500; c1000 = effect of congruent primes in trial 1000; c1500 = effect of congruent primes in trial 1500; i500 = effect of incongruent primes in trial 500; i1000 = effect of incongruent primes in trial 1000; i1500 = effect of incongruent primes in trial 1500.

598

599 Based on Figure 8 and Table 5, we see that throughout the experiment (trials 500,
600 1000, and 1500), congruent primes have a positive effect on logit-ca(t) in time bins
601 (200,240], (240,280], (280,320], and (320,360], relative to the logit-ca(t) estimates in the
602 baseline condition (blank prime; red striped lines in Figure 8). For example, the odds ratio
603 for congruent primes in bin (320,360] in trial 500 shows that the odds of a correct response
604 are estimated to be 8.89 times higher than the odds of a correct response when there is no
605 prime. Incongruent primes have a negative effect on logit-ca(t) in time bins (200,240],
606 (240,280], and (280,320] throughout the experiment, relative to the logit-ca(t) estimates in
607 the baseline condition (no prime; red striped lines).

608 To conclude this Tutorial 2b, Figure 9 shows the model-based ca(t) functions for each
609 prime type for participant 6, in trial 500, 1000, and 1500.

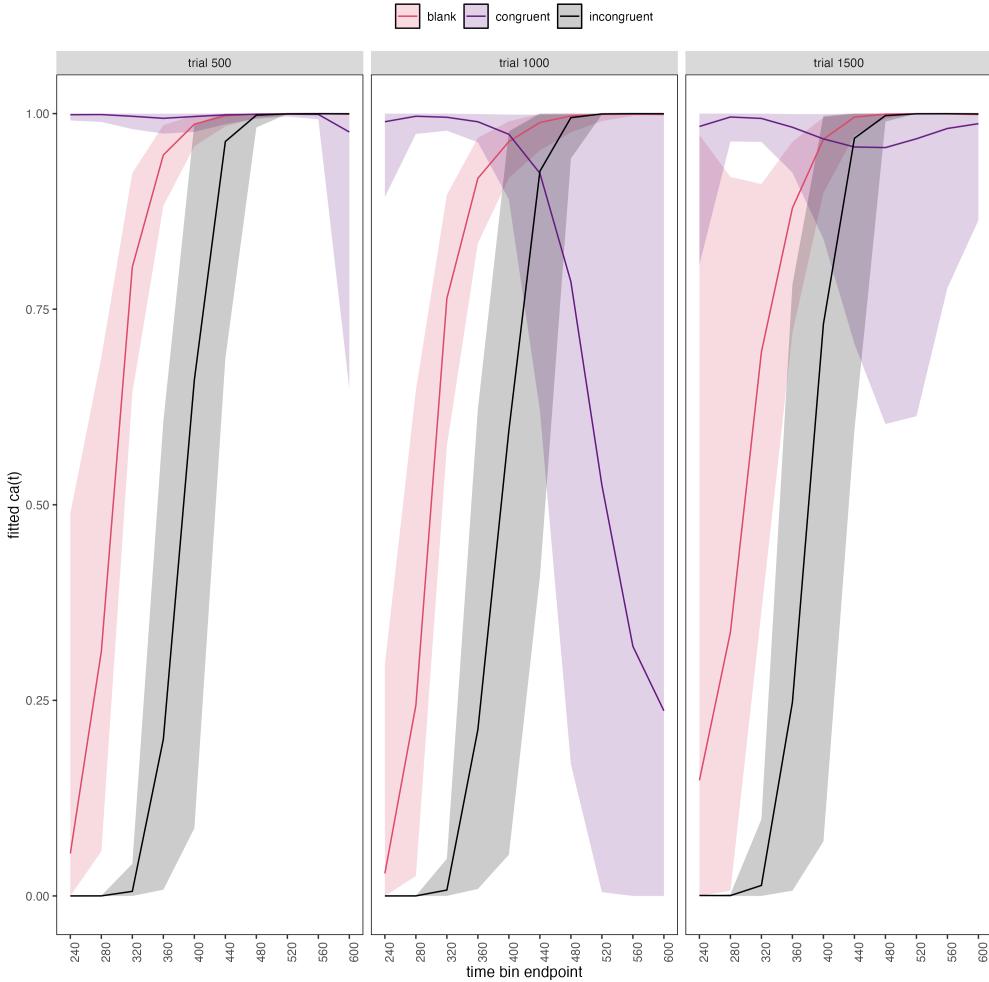


Figure 9. Model-based $ca(t)$ functions for participant 6 in trial 500, 1000, and 1500.

610 4.5 Tutorial 3a: Fitting Frequentist hazard models

611 In this fifth tutorial we illustrate how to fit a multilevel regression model to RT data
 612 in the frequentist framework, for the data set used in Tutorial 1a. The general process is
 613 similar to that in Tutorial 2a, except that there are no priors to set. For illustration
 614 purposes, we only fitted the effects from model M3 (see Tutorial 2a) using the function
 615 `glmer()` from the R package `lme4`. Alternatively, one could also use the function
 616 `glmmPQL()` from the R package `MASS` (Ripley et al., 2024). The resulting hazard model
 617 is called `M3_f`.

618 In Figure 10 we compare the parameter estimates of model M3 from `brm()` with those

619 of model M3_f from `glmer()`.

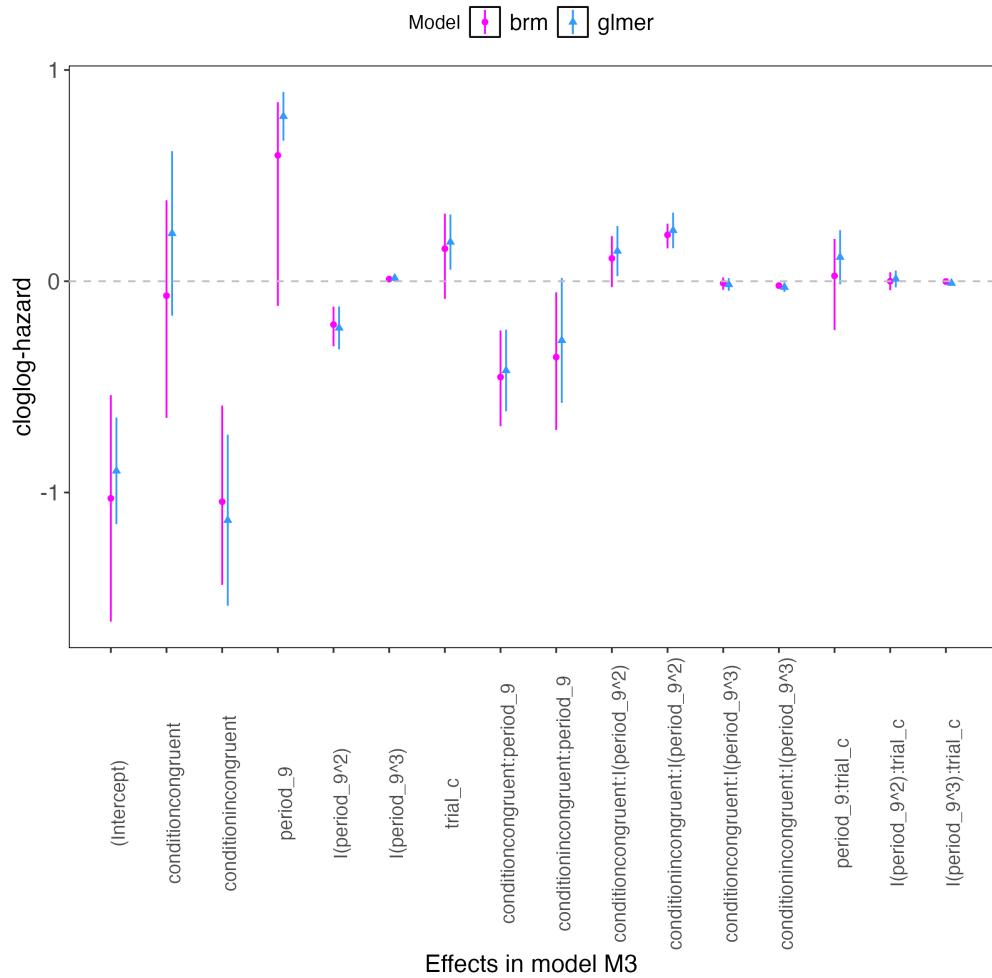


Figure 10. Parameter estimates for model M3 from `brm()` – means and 95% quantile intervals – and model M3_f from `glmer()` – maximum likelihood estimates and 95% confidence intervals.

620 Figure 10 confirms that the parameter estimates from both Bayesian and frequentist

621 models are pretty similar. However, the random effects structure of model M3_f was

622 already too complex for the frequentist model as it did not converge and resulted in a

623 singular fit. This is of course one of the reasons why Bayesian modeling has become so

624 popular in recent years. But the price you pay for being able to fit more complex random

625 effects models in a Bayesian framework is computation time. In other words, as we have
626 noted throughout, some of the Bayesian models in Tutorials 2a and 2b took several hours
627 to build.

628 **4.6 Tutorial 3b: Fitting Frequentist conditional accuracy models**

629 In this sixth tutorial we illustrate how to fit a multilevel regression model to the
630 timed accuracy data in the frequentist framework, for the data set used in Tutorial 1a. For
631 illustration purposes, we only fitted the effects from model M3 (see Tutorial 2a) using the
632 function `glmer()` from the R package `lme4`. Alternatively, one could also use the function
633 `glmmPQL()` from the R package `MASS` (Ripley et al., 2024). Again, the resulting
634 conditional accuracy model `M3_ca_f` did not converge and resulted in a singular fit.

635 **5. Discussion**

636 This main motivation for writing this paper is the observation that event history and
637 SAT analyses remain under-used in psychological research, which means the field of
638 psychological research is not taking full advantage of the many benefits EHA/SAT provides
639 compared to more conventional analyses. By providing a freely available set of tutorials,
640 which provide step-by-step guidelines and ready-to-use R code, we hope that researchers
641 will feel more comfortable using EHA/SAT in the future. Indeed, we hope that our
642 tutorials may help to overcome a barrier to entry with EHA/SAT, which is the increase in
643 analytical complexity compared to mean-average comparisons. While we have focused here
644 on within-subject, factorial, small- N designs, it is important to realize that EHA/SAT can
645 be applied to other designs as well (large- N designs with only one measurement per
646 subject, between-subject designs, etc.). As such, the general workflow and associated code
647 can be modified and applied more broadly to other contexts and research questions. In the
648 following, we discuss issues relating to model complexity versus interpretability, individual
649 differences, limitations of the approach, and future extensions.

650 **5.1 Advantages of hazard analysis**

651 Statisticians and mathematical psychologists recommend focusing on the hazard
652 function when analyzing time-to-event data for various reasons. First, as discussed by
653 Holden, Van Orden, and Turvey (2009), “probability density [and mass] functions can
654 appear nearly identical, both statistically and to the naked eye, and yet are clearly different
655 on the basis of their hazard functions (but not vice versa). Hazard functions are thus more
656 diagnostic than density functions” (p. 331) when one is interested in studying the detailed
657 shape of a RT distribution (see also Figure 1 in Panis, Schmidt, et al., 2020). Therefore,
658 when the goal is to study how psychological effects change over time, hazard and
659 conditional accuracy functions are preferred.

660 Second, because RT distributions may differ from one another in multiple ways,
661 Townsend (1990) developed a dominance hierarchy of statistical differences between two
662 arbitrary distributions A and B. For example, if $h_A(t) > h_B(t)$ for all t, then both hazard
663 functions are said to show a complete ordering. Townsend (1990) concluded that stronger
664 conclusions can be drawn from data when comparing the hazard functions using EHA. For
665 example, when mean A < mean B, the hazard functions might show a complete ordering
666 (i.e., for all t), a partial ordering (e.g., only for $t > 300$ ms, or only for $t < 500$ ms), or they
667 may cross each other one or more times.

668 Third, EHA does not discard right-censored observations when estimating hazard
669 functions, that is, trials for which we do not observe a response during the data collection
670 period in a trial so that we only know that the RT must be larger than some value (e.g.,
671 the response deadline). This is important because although a few right-censored
672 observations are inevitable in most RT tasks, a lot of right-censored observations are
673 expected in experiments on masking, the attentional blink, and so forth. In other words, by
674 using EHA you can analyze RT data from experiments that typically do not measure
675 response times. As a result, EHA can also deal with long RTs in experiments without a

676 response deadline, which are typically treated as outliers and are discarded before
677 calculating a mean. This orthodox procedure leads to underestimation of the true mean.
678 By introducing a fixed censoring time for all trials at the end of the analysis time window,
679 trials with long RTs are not discarded but contribute to the risk set of each bin.

680 Fourth, hazard modeling allows incorporating time-varying explanatory covariates
681 such as heart rate, electroencephalogram (EEG) signal amplitude, gaze location, etc.
682 (Allison, 2010). This is useful for linking physiological effects to behavioral effects when
683 performing cognitive psychophysiology (Meyer, Osman, Irwin, & Yantis, 1988).

684 Finally, as explained by Kelso, Dumas, and Tognoli (2013), it is crucial to first have a
685 precise description of the macroscopic behavior of a system (here: $h(t)$ and possibly $ca(t)$
686 functions) in order to know what to derive on the microscopic level. EHA can thus solve
687 the problem of model mimicry, i.e., the fact that different computational models can often
688 predict the same mean RTs as observed in the empirical data, but not necessarily the
689 detailed shapes of the empirical RT hazard distributions. Also, fitting parametric functions
690 or computational models to data without studying the shape of the empirical discrete-time
691 $h(t)$ and $ca(t)$ functions can miss important features in the data (Panis, Moran, et al.,
692 2020; Panis & Schmidt, 2016).

693 5.2 Model complexity versus interpretability

694 Models for discrete-time $h(t)$ and $ca(t)$ can quickly become very complex when adding
695 more than 1 time scale, due to the many possible higher-order interactions. For example,
696 model M4 contains two time scales as covariates: the passage of time on the across-bin or
697 within-trial time scale (variable period_9), and the passage of time on the across-trial or
698 within-experiment time scale (variable trial_c). However, when trials are presented in
699 blocks, and blocks of trials within sessions, and when the experiment comprises three
700 sessions, then four time scales can be defined (across-bin or within-trial, across-trial or

701 within-block, across-block or within-session, and across-session or within-experiment).
702 From a theoretical perspective, adding more than 1 time scale – and their interactions – is
703 important to capture plasticity and other learning effects (e.g., proactive control) that play
704 out on such longer time scales (across-trials, across-blocks, across-sessions), and that are
705 probably present in each experiment in general. From a practical perspective, therefore, it
706 might be interesting for interpretational purposes to limit the number of experimental
707 predictor variables, because adding time scales quickly increases model complexity.

708 **5.3 Individual differences**

709 One important issue is that of possible individual differences in the overall location of
710 the distribution, and the time course of psychological effects. For example, when you wait
711 for a response of the participant on each trial, you allow the participant to have control
712 over the trial duration, and some participants might respond only when they are confident
713 that their emitted response will be correct. These issues can be avoided by introducing a
714 (relatively short) response deadline in each trial, e.g., 600 ms for simple detection tasks,
715 1000 ms for more difficult discrimination tasks, or 2 s for tasks requiring extended high-level
716 processing. Because EHA can deal in a straightforward fashion with right-censored
717 observations (i.e., trials without an observed response), introducing a response deadline is
718 recommended when designing RT experiments. Furthermore, introducing a response
719 deadline and asking participants to respond before the deadline as much as possible, will
720 also lead to individual distributions that overlap in time, which is important when selecting
721 a common analysis time window when fitting hazard and conditional accuracy models.

722 But even when using a response deadline, participants can differ qualitatively in the
723 effects they display (see Panis, 2020). One way to deal with this is to describe and
724 interpret the different patterns. Another way is to run a clustering algorithm on the
725 individual hazard estimates across all conditions. The obtained dendrogram can then be
726 used to identify a (hopefully big) cluster of participants that behave similarly, and to

727 identify a (hopefully small) cluster of participants with outlying behavioral patterns. One
728 might then exclude the outlying participants before fitting a hazard model.

729 Another approach to deal with individual differences is Bayesian prevalence (Ince,
730 Paton, Kay, & Schyns, 2021). This method looks at effects within each individual in the
731 study and asks how likely it would be to see the same result if the experiment was repeated
732 with a new person chosen from the wider population at random. This approach allows one
733 to quantify how typical or uncommon an observed effect is in the population, and the
734 uncertainty around this estimate.

735 **5.4 Limitations**

736 Compared to the orthodox method – comparing mean-averages between conditions –
737 the most important limitation of multi-level hazard and conditional accuracy modeling is
738 that it might take a long time to estimate the parameters using Bayesian methods or the
739 model might have to be simplified significantly to use frequentist methods.

740 Another issue is that you need a relatively large number of trials per condition to
741 estimate the hazard function with high temporal resolution, which is required when testing
742 predictions of process models of cognition. Indeed, in general, there is a trade-off between
743 the number of trials per condition and the temporal resolution (i.e., bin width) of the
744 hazard function. Therefore, we recommend researchers to collect as many trials as possible
745 per experimental condition, given the available resources and considering the participant
746 experience (e.g., fatigue and boredom). For instance, if the maximum session length
747 deemed reasonable is between 1 and 2 hours, what is the maximum number of trials per
748 condition that you could reasonably collect? After consideration, it might be worth
749 conducting multiple testing sessions per participant and/or reducing the number of
750 experimental conditions. Finally, there is a user-friendly online tool for calculating
751 statistical power as a function of the number of trials as well as the number of participants,

752 and this might be worth consulting to guide the research design process (Baker et al., 2021).

753 We did not discuss continuous-time EHA, nor continuous-time SAT analysis. As
754 indicated by Allison (2010), learning discrete-time EHA methods first will help in learning
755 continuous-time methods. Given that RT is typically treated as a continuous variable, it is
756 possible that continuous-time methods will ultimately prevail. However, they require much
757 more data to estimate the continuous-time hazard (rate) function well. Thus, by trading a
758 bit of temporal resolution for a lower number of trials, discrete-time methods seem ideal for
759 dealing with typical psychological time-to-event data sets for which there are less than
760 ~200 trials per condition per experiment.

761 **5.5 Extensions**

762 The hazard models in this tutorial assume that there is one event of interest. For RT
763 data, this event constitutes a single transition between an “idle” state and a “responded”
764 state. However, in certain situations, more than one event of interest might exist. For
765 example, in a medical or health-related context, an individual might transition back and
766 forth between a “healthy” state and a “depressed” state, before being absorbed into a final
767 “death” state. When you have data on the timing of these transitions, one can apply
768 multi-state hazard models, which generalize event history analysis to transitions between
769 three or more states (Steele, Goldstein, & Browne, 2004). Also, the predictor variables in
770 this tutorial are time-invariant, i.e., their value did not change over the course of a trial.
771 Thus, another extension is to include time-varying predictors, i.e., predictors whose value
772 can change across the time bins within a trial (Allison, 2010). For example, when gaze
773 position is tracked during a visual search trial, the gaze-target distance will vary during a
774 trial when the eyes move around before a manual response is given; shorter gaze-target
775 distances should be associated with a higher hazard of response occurrence. Note that the
776 effect of a time-varying predictor (e.g., an occipital EEG signal) can itself vary over time.

777

6. Conclusions

778 RT and accuracy distributions are a rich source of information on the time course of
779 cognitive processing, which have been largely undervalued in the history of experimental
780 psychology and cognitive neuroscience. Statistically controlling for the passage of time
781 during data analysis is equally important as experimental control during the design of an
782 experiment, to better understand human behavior in experimental paradigms. We hope
783 that by providing a set of hands-on, step-by-step tutorials, which come with custom-built
784 and freely available code, researchers will feel more comfortable embracing event history
785 analysis and investigating the temporal profile of cognitive states. On a broader level, we
786 think that wider adoption of such approaches will have a meaningful impact on the
787 inferences drawn from data, as well as the development of theories regarding the structure
788 of cognition.

789

References

- 790 Allison, P. D. (1982). Discrete-Time Methods for the Analysis of Event Histories.
791 *Sociological Methodology*, 13, 61. <https://doi.org/10.2307/270718>
- 792 Allison, P. D. (2010). *Survival analysis using SAS: A practical guide* (2. ed). Cary, NC:
793 SAS Press.
- 794 Aust, F. (2019). *Citr: 'RStudio' add-in to insert markdown citations*. Retrieved from
795 <https://github.com/crsh/citr>
- 796 Aust, F., & Barth, M. (2023). *papaja: Prepare reproducible APA journal articles with R*
797 *Markdown*. Retrieved from <https://github.com/crsh/papaja>
- 798 Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., &
799 Andrews, T. J. (2021). Power contours: Optimising sample size and precision in
800 experimental psychology and human neuroscience. *Psychological Methods*, 26(3),
801 295–314. <https://doi.org/10.1037/met0000337>
- 802 Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for
803 confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*,
804 68(3), 10.1016/j.jml.2012.11.001. <https://doi.org/10.1016/j.jml.2012.11.001>
- 805 Barth, M. (2023). *tinylabes: Lightweight variable labels*. Retrieved from
806 <https://cran.r-project.org/package=tinylabes>
- 807 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects
808 models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
809 <https://doi.org/10.18637/jss.v067.i01>
- 810 Bates, D., Maechler, M., & Jagan, M. (2024). *Matrix: Sparse and dense matrix classes and*
811 *methods*. Retrieved from <https://CRAN.R-project.org/package=Matrix>
- 812 Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in r
813 using futures. *The R Journal*, 13(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>
- 814 Blossfeld, H.-P., & Rohwer, G. (2002). *Techniques of event history modeling: New*
815 *approaches to causal analysis*, 2nd ed (pp. x, 310). Mahwah, NJ, US: Lawrence

- 816 Erlbaum Associates Publishers.
- 817 Box-Steffensmeier, J. M. (2004). Event history modeling: A guide for social scientists.
818 Cambridge: University Press.
- 819 Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan.
820 *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- 821 Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms.
822 *The R Journal*, 10(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>
- 823 Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal*
824 *of Statistical Software*, 100(5), 1–54. <https://doi.org/10.18637/jss.v100.i05>
- 825 Eddelbuettel, D., & Balamuta, J. J. (2018). Extending R with C++: A Brief Introduction
826 to Rcpp. *The American Statistician*, 72(1), 28–36.
827 <https://doi.org/10.1080/00031305.2017.1375990>
- 828 Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal*
829 *of Statistical Software*, 40(8), 1–18. <https://doi.org/10.18637/jss.v040.i08>
- 830 Gabry, J., Češnovar, R., Johnson, A., & Broder, S. (2024). *Cmdstanr: R interface to*
831 *'CmdStan'*. Retrieved from <https://github.com/stan-dev/cmdstanr>
- 832 Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization
833 in bayesian workflow. *J. R. Stat. Soc. A*, 182, 389–402.
834 <https://doi.org/10.1111/rssc.12378>
- 835 Girard, J. (2024). *Standist: What the package does (one line, title case)*. Retrieved from
836 <https://github.com/jmgirard/standist>
- 837 Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate.
838 *Journal of Statistical Software*, 40(3), 1–25. Retrieved from
839 <https://www.jstatsoft.org/v40/i03/>
- 840 Holden, J. G., Van Orden, G. C., & Turvey, M. T. (2009). Dispersion of response times
841 reveals cognitive dynamics. *Psychological Review*, 116(2), 318–342.
842 <https://doi.org/10.1037/a0014849>

- 843 Hosmer, D. W., Lemeshow, S., & May, S. (2011). *Applied Survival Analysis: Regression*
844 *Modeling of Time to Event Data* (2nd ed). Hoboken: John Wiley & Sons.
- 845 Ince, R. A., Paton, A. T., Kay, J. W., & Schyns, P. G. (2021). Bayesian inference of
846 population prevalence. *eLife*, 10, e62461. <https://doi.org/10.7554/eLife.62461>
- 847 Kantowitz, B. H., & Pachella, R. G. (2021). The Interpretation of Reaction Time in
848 Information-Processing Research 1. *Human Information Processing*, 41–82.
849 <https://doi.org/10.4324/9781003176688-2>
- 850 Kay, M. (2023). *tidybayes: Tidy data and geoms for Bayesian models*.
851 <https://doi.org/10.5281/zenodo.1308151>
- 852 Kelso, J. A. S., Dumas, G., & Tognoli, E. (2013). Outline of a general theory of behavior
853 and brain coordination. *Neural Networks: The Official Journal of the International*
854 *Neural Network Society*, 37, 120–131. <https://doi.org/10.1016/j.neunet.2012.09.003>
- 855 Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing,
856 estimation, meta-analysis, and power analysis from a Bayesian perspective.
857 *Psychonomic Bulletin & Review*, 25(1), 178–206.
858 <https://doi.org/10.3758/s13423-016-1221-4>
- 859 Kurz, A. S. (2023a). *Applied longitudinal data analysis in brms and the tidyverse* (version
860 0.0.3). Retrieved from <https://bookdown.org/content/4253/>
- 861 Kurz, A. S. (2023b). *Statistical rethinking with brms, ggplot2, and the tidyverse: Second*
862 *edition* (version 0.4.0). Retrieved from <https://bookdown.org/content/4857/>
- 863 Landes, J., Engelhardt, S. C., & Pelletier, F. (2020). An introduction to event history
864 analyses for ecologists. *Ecosphere*, 11(10), e03238. <https://doi.org/10.1002/ecs2.3238>
- 865 Luce, R. D. (1991). *Response times: Their role in inferring elementary mental organization*
866 (1. issued as paperback). Oxford: Univ. Press.
- 867 McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and*
868 *Stan* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315372495>
- 869 Meyer, D. E., Osman, A. M., Irwin, D. E., & Yantis, S. (1988). Modern mental

- 870 chronometry. *Biological Psychology*, 26(1-3), 3–67.
- 871 [https://doi.org/10.1016/0301-0511\(88\)90013-0](https://doi.org/10.1016/0301-0511(88)90013-0)
- 872 Müller, K., & Wickham, H. (2023). *Tibble: Simple data frames*. Retrieved from
873 <https://CRAN.R-project.org/package=tibble>
- 874 Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes*. Retrieved from
875 <https://CRAN.R-project.org/package=RColorBrewer>
- 876 Panis, S. (2020). How can we learn what attention is? Response gating via multiple direct
877 routes kept in check by inhibitory control processes. *Open Psychology*, 2(1), 238–279.
878 <https://doi.org/10.1515/psych-2020-0107>
- 879 Panis, S., Moran, R., Wolkersdorfer, M. P., & Schmidt, T. (2020). Studying the dynamics
880 of visual search behavior using RT hazard and micro-level speed–accuracy tradeoff
881 functions: A role for recurrent object recognition and cognitive control processes.
882 *Attention, Perception, & Psychophysics*, 82(2), 689–714.
883 <https://doi.org/10.3758/s13414-019-01897-z>
- 884 Panis, S., Schmidt, F., Wolkersdorfer, M. P., & Schmidt, T. (2020). Analyzing Response
885 Times and Other Types of Time-to-Event Data Using Event History Analysis: A Tool
886 for Mental Chronometry and Cognitive Psychophysiology. *I-Perception*, 11(6),
887 2041669520978673. <https://doi.org/10.1177/2041669520978673>
- 888 Panis, S., & Schmidt, T. (2016). What Is Shaping RT and Accuracy Distributions? Active
889 and Selective Response Inhibition Causes the Negative Compatibility Effect. *Journal of*
890 *Cognitive Neuroscience*, 28(11), 1651–1671. https://doi.org/10.1162/jocn_a_00998
- 891 Panis, S., & Schmidt, T. (2022). When does “inhibition of return” occur in spatial cueing
892 tasks? Temporally disentangling multiple cue-triggered effects using response history
893 and conditional accuracy analyses. *Open Psychology*, 4(1), 84–114.
894 <https://doi.org/10.1515/psych-2022-0005>
- 895 Panis, S., Torfs, K., Gillebert, C. R., Wageman, J., & Humphreys, G. W. (2017).
896 Neuropsychological evidence for the temporal dynamics of category-specific naming.

- 897 *Visual Cognition*, 25(1-3), 79–99. <https://doi.org/10.1080/13506285.2017.1330790>
- 898 Panis, S., & Wagemans, J. (2009). Time-course contingencies in perceptual organization
899 and identification of fragmented object outlines. *Journal of Experimental Psychology:
900 Human Perception and Performance*, 35(3), 661–687.
901 <https://doi.org/10.1037/a0013547>
- 902 Pedersen, T. L. (2024). *Patchwork: The composer of plots*. Retrieved from
903 <https://CRAN.R-project.org/package=patchwork>
- 904 Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in s and s-PLUS*. New York:
905 Springer. <https://doi.org/10.1007/b98882>
- 906 R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna,
907 Austria: R Foundation for Statistical Computing. Retrieved from
908 <https://www.R-project.org/>
- 909 Ripley, B., Venables, B., Bates, D. M., ca 1998), K. H. (partial. port, ca 1998), A. G.
910 (partial. port, & polr), D. F. (support. functions for. (2024). *MASS: Support Functions
911 and Datasets for Venables and Ripley's MASS*.
- 912 Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling
913 Change and Event Occurrence*. Oxford, New York: Oxford University Press.
- 914 Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design.
915 *Psychonomic Bulletin & Review*, 25(6), 2083–2101.
916 <https://doi.org/10.3758/s13423-018-1451-8>
- 917 Steele, F., Goldstein, H., & Browne, W. (2004). A general multilevel multistate competing
918 risks model for event history data, with an application to a study of contraceptive use
919 dynamics. *Statistical Modelling*, 4(2), 145–159.
920 <https://doi.org/10.1191/1471082X04st069oa>
- 921 Teachman, J. D. (1983). Analyzing social processes: Life tables and proportional hazards
922 models. *Social Science Research*, 12(3), 263–301.
923 [https://doi.org/10.1016/0049-089X\(83\)90015-7](https://doi.org/10.1016/0049-089X(83)90015-7)

- 924 Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical
925 distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, 108(3),
926 551–567. <https://doi.org/10.1037/0033-2909.108.3.551>
- 927 Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics.
928 *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)
- 929 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New
930 York. Retrieved from <https://ggplot2.tidyverse.org>
- 931 Wickham, H. (2023a). *Forcats: Tools for working with categorical variables (factors)*.
932 Retrieved from <https://forcats.tidyverse.org/>
- 933 Wickham, H. (2023b). *Stringr: Simple, consistent wrappers for common string operations*.
934 Retrieved from <https://stringr.tidyverse.org>
- 935 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ...
936 Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43),
937 1686. <https://doi.org/10.21105/joss.01686>
- 938 Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). *R for data science: Import,
939 tidy, transform, visualize, and model data* (2nd edition). Beijing Boston Farnham
940 Sebastopol Tokyo: O'Reilly.
- 941 Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A
942 grammar of data manipulation*. Retrieved from <https://dplyr.tidyverse.org>
- 943 Wickham, H., & Henry, L. (2023). *Purrr: Functional programming tools*. Retrieved from
944 [https://purrr.tidyverse.org/](https://purrr.tidyverse.org)
- 945 Wickham, H., Hester, J., & Bryan, J. (2024). *Readr: Read rectangular text data*. Retrieved
946 from <https://readr.tidyverse.org>
- 947 Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data*. Retrieved from
948 <https://tidyr.tidyverse.org>
- 949 William Matthew Makeham. (1860). *On the Law of Mortality and the Construction of
950 Annuity Tables*. The Assurance Magazine, and Journal of the Institute of Actuaries.

- 951 Wolkersdorfer, M. P., Panis, S., & Schmidt, T. (2020). Temporal dynamics of sequential
952 motor activation in a dual-prime paradigm: Insights from conditional accuracy and
953 hazard functions. *Attention, Perception, & Psychophysics*, 82(5), 2581–2602.
954 <https://doi.org/10.3758/s13414-020-02010-5>

955

Supplementary material

956 **A. Definitions of discrete-time hazard, survivor, probability mass, and**
 957 **conditional accuracy functions**

958 The shape of a distribution of waiting times can be described in multiple ways (Luce,
 959 1991). After dividing time in discrete, contiguous time bins indexed by t , let RT be a
 960 discrete random variable denoting the rank of the time bin in which a particular person's
 961 response occurs in a particular experimental condition. Because waiting times can only
 962 increase, discrete-time EHA focuses on the discrete-time hazard function

963
$$h(t) = P(RT = t | RT \geq t) \quad (1)$$

964 and the discrete-time survivor function

965
$$S(t) = P(RT > t) = [1-h(t)].[1-h(t-1)].[1-h(t-2)] \dots [1-h(1)] \quad (2)$$

966 and not on the probability mass function

967
$$P(t) = P(RT = t) = h(t).S(t-1) \quad (3)$$

968 nor the cumulative distribution function

969
$$F(t) = P(RT \leq t) = 1-S(t) \quad (4)$$

970 The discrete-time hazard function of event occurrence gives you for each bin the
 971 probability that the event occurs (sometime) in that bin, given that the event has not
 972 occurred yet in previous bins. While the discrete-time hazard function assesses the unique
 973 risk of event occurrence associated with each time bin, the discrete-time survivor function
 974 cumulates the bin-by-bin risks of event *nonoccurrence* to obtain the probability that the
 975 event occurs after bin t . The probability mass function cumulates the risk of event
 976 occurrence in bin t with the risks of event nonoccurrence in bins 1 to $t-1$. From equation 3
 977 we find that hazard in bin t is equal to $P(t)/S(t-1)$.

978 For two-choice RT data, the discrete-time hazard function can be extended with the

979 discrete-time conditional accuracy function

980 $ca(t) = P(\text{correct} \mid RT = t)$ (5)

981 which gives you for each bin the probability that a response is correct given that it is
 982 emitted in time bin t (Allison, 2010; Kantowitz & Pachella, 2021; Wickelgren, 1977). This
 983 latter function is also known as the micro-level speed-accuracy tradeoff (SAT) function.

984 The survivor function provides a context for the hazard function, as $S(t-1) = P(RT >$
 985 $t-1) = P(RT \geq t)$ tells you on how many percent of the trials the estimate $h(t) = P(RT =$
 986 $t \mid RT \geq t)$ is based. The probability mass function provides a context for the conditional
 987 accuracy function, as $P(t) = P(RT = t)$ tells you on how many percent of the trials the
 988 estimate $ca(t) = P(\text{correct} \mid RT = t)$ is based.

989 When time is treated as a continuous variable, let RT be a continuous random variable
 990 denoting a particular person's response time in a particular experimental condition.

991 Continuous-time EHA does not focus on the cumulative distribution function $F(t) = P(RT$
 992 $\leq t)$ and its derivative, the probability density function $f(t) = F(t)'$, but on the survivor
 993 function $S(t) = P(RT > t)$ and the hazard rate function $\lambda(t) = f(t)/S(t)$. The hazard rate
 994 function gives you the instantaneous *rate* of event occurrence at time point t , given that
 995 the event has not occurred yet.

996 B. Custom functions for descriptive discrete-time hazard analysis

997 We defined 13 custom functions that we list here.

- 998 • `censor(df,timeout,bin_width)` : divide the time segment $(0,timeout]$ in bins, identify
 999 any right-censored observations, and determine the discrete RT (time bin rank)
- 1000 • `ptb(df)` : transform the person-trial data set to the person-trial-bin data set
- 1001 • `setup_lt(ptb)` : set up a life table for each level of 1 independent variable

- 1002 • setup_lt_2IV(ptb) : set up a life table for each combination of levels of 2
1003 independent variables
- 1004 • calc_ca(df) : estimate the conditinal accuracies when there is 1 independent variable
- 1005 • calc_ca_2IV(df) : estimate the conditional accuraiies when there are 2 independent
1006 variables
- 1007 • join_lt_ca(df1,df2) : add the ca(t) estimates to the life tables (1 independent
1008 variable)
- 1009 • join_lt_ca_2IV(df1,df2) : add the ca(t) estimates to the life tables (2 independent
1010 variables)
- 1011 • extract_median(df) : estimate quantiles $S(t)._{50}$ (1 independent variable)
- 1012 • extract_median_2IV(df) : estimate quantiles $S(t)._{50}$ (2 independent variables)
- 1013 • plot_oha(df,subj,haz_yaxis) : create plots of the discrete-time functions (1
1014 independent variable)
- 1015 • plot_oha_2IV(df,subj,haz_yaxis) : create plots of the discrete-time functions (2
1016 independent variables)
- 1017 • plot_oha_agg(df,subj,haz_yaxis) : create 1 plot for data aggregated across
1018 participants (1 independent variable)

1019 When you want to analyse simple RT data from a detection experiment with one
1020 independent variable, the functions calc_ca() and join_lt_ca() should not be used, and
1021 the code to plot the conditional accuracy functions should be removed from the function
1022 plot_oha(). When you want to analyse simple RT data from a detection experiment with
1023 two independent variables, the functions calc_ca_2IV() and join_lt_ca_2IV() should not
1024 be used, and the code to plot the conditional accuracy functions should be removed from
1025 the function plot_oha_2IV().

1026 **C. Link functions**

1027 Popular link functions include the logit link and the complementary log-log link, as
 1028 shown in Figure 11.

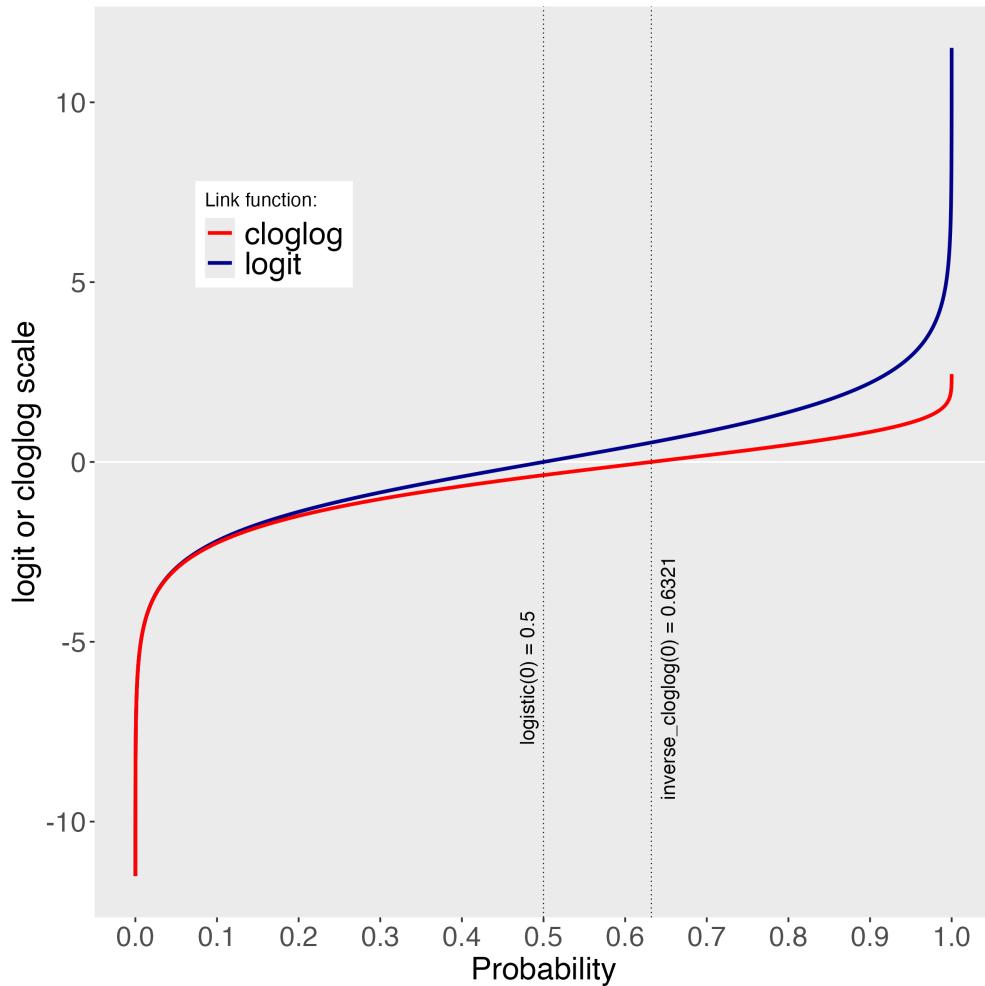


Figure 11. The logit and cloglog link functions.

1029 **D. Regression equations**

1030 An example (single-level) discrete-time hazard model with three predictors (TIME,
 1031 X₁, X₂), the cloglog link function, and a third-order polynomial specification for TIME can
 1032 be written as follows:

$$\begin{aligned} 1033 \quad & \text{cloglog}[h(t)] = \ln(-\ln[1-h(t)]) = [\beta_0 \text{ONE} + \beta_1(\text{TIME}-9) + \beta_2(\text{TIME}-9)^2] + [\beta_3 X_1 + \beta_4 X_2 \\ 1034 \quad & + \beta_5 X_2(\text{TIME}-9)] \end{aligned} \quad (6)$$

1035 The main predictor variable TIME is the time bin index t that is centered on value 9
 1036 in this example. The first set of terms within brackets, the parameters β_0 to β_2 multiplied
 1037 by their polynomial specifications of (centered) time, represents the shape of the baseline
 1038 cloglog-hazard function (i.e., when all predictors X_i take on a value of zero). The second
 1039 set of terms (the beta parameters β_3 to β_5) represents the vertical shift in the baseline
 1040 cloglog-hazard for a 1 unit increase in the respective predictor variable. Predictors can be
 1041 discrete, continuous, and time-varying or time-invariant. For example, the effect of a 1 unit
 1042 increase in X_1 is to vertically shift the whole baseline cloglog-hazard function by β_1
 1043 cloglog-hazard units. However, if the predictor interacts linearly with TIME (see X_2 in the
 1044 example), then the effect of a 1 unit increase in X_2 is to vertically shift the predicted
 1045 cloglog-hazard in bin 9 by β_2 cloglog-hazard units (when $\text{TIME}-9 = 0$), in bin 10 by $\beta_2 + \beta_3$ cloglog-hazard units (when $\text{TIME}-9 = 1$), and so forth. To interpret the effects of a
 1046 predictor, its β parameter is exponentiated, resulting in a hazard ratio (due to the use of
 1047 the cloglog link). When using the logit link, exponentiating a β parameter results in an
 1048 odds ratio.
 1049

1050 An example (single-level) discrete-time hazard model with a general specification for
 1051 TIME (separate intercepts for each of six bins, where D1 to D6 are binary variables
 1052 identifying each bin) and a single predictor (X_1) can be written as follows:

$$1053 \quad \text{cloglog}[h(t)] = [\beta_0 D1 + \beta_1 D2 + \beta_2 D3 + \beta_3 D4 + \beta_4 D5 + \beta_5 D6] + [\beta_6 X_1] \quad (7)$$

1054 **E. Prior distributions**

1055 To gain a sense of what prior *logit* values would approximate a uniform distribution
 1056 on the probability scale, Kurz (2023a) simulated a large number of draws from the
 1057 Uniform(0,1) distribution, converted those draws to the log-odds metric, and fitted a

1058 Student's t distribution. Row C in Figure 12 shows that using a t-distribution with 7.61
 1059 degrees of freedom and a scale parameter of 1.57 as a prior on the logit scale, approximates
 1060 a uniform distribution on the probability scale. According to Kurz (2023a), such a prior
 1061 might be a good prior for the intercept(s) in a logit-hazard model, while the $N(0,1)$ prior in
 1062 row D might be a good prior for the non-intercept parameters in a logit-hazard model, as it
 1063 gently regularizes p towards .5 (i.e., a zero effect on the logit scale).

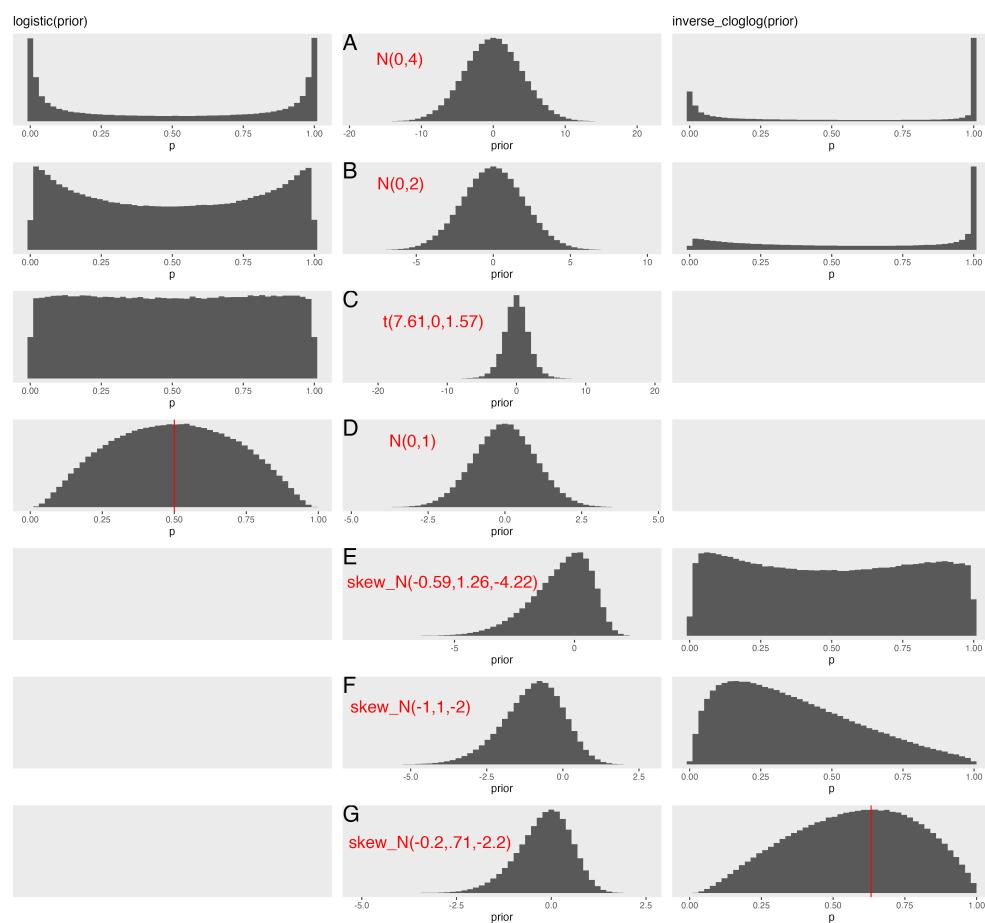


Figure 12. Prior distributions on the logit and/or cloglog scales (middle column), and their implications on the probability scale after applying the inverse-logit (or logistic) transformation (left column), and the inverse-cloglog transformation (right column).

1064 To gain a sense of what prior *cloglog* values would approximate a uniform distribution
 1065 on the hazard probability scale, we followed Kurz's approach and simulated a large number

of draws from the Uniform(0,1) distribution, converted them to the cloglog metric, and fitted a skew-normal model (due to the asymmetry of the cloglog link function). Row E shows that using a skew-normal distribution with a mean of -0.59, a standard deviation of 1.26, and a skewness of -4.22 as a prior on the cloglog scale, approximates a uniform distribution on the probability scale. However, because hazard values below .5 are more likely in RT studies, using a skew-normal distribution with a mean of -1, a standard deviation of 1, and a skewness of -2 as a prior on the cloglog scale (row F), might be a good weakly informative prior for the intercept(s) in a cloglog-hazard model. A skew-normal distribution with a mean of -0.2, a standard deviation of 0.71, and a skewness of -2.2 might be a good weakly informative prior for the non-intercept parameters in a cloglog-hazard model as it gently regularizes p towards .6321 (i.e., a zero effect on the cloglog scale).