

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.

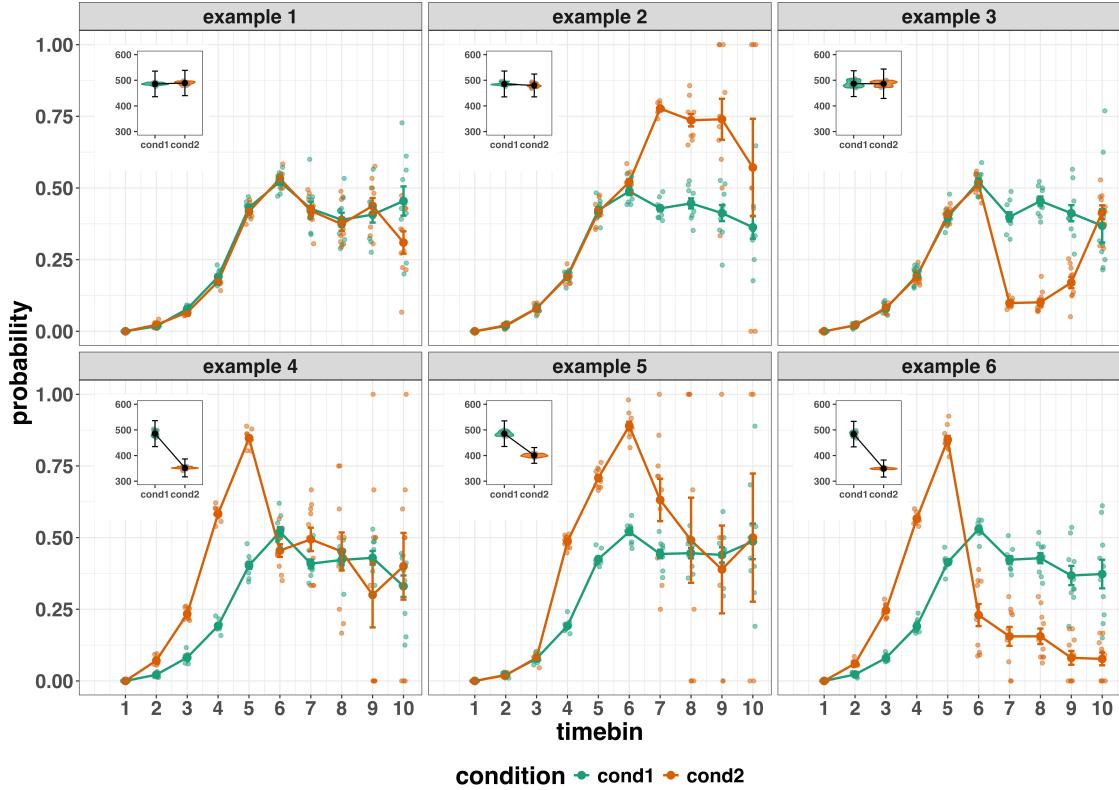


Figure 1. Means versus distributional shapes for six different simulated data set examples. The first second after stimulus onset is divided in ten bins of 100 ms. For our purposes here, it is enough to know that the distributions plotted represent the probability of an event occurring in that timebin, 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.

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

107 In Tutorial 1a, we illustrate how to process or “wrangle” a previously published RT +
108 accuracy data set to calculate descriptive statistics when there is one independent variable.
109 The descriptive statistics are plotted, and we comment on their interpretation. In Tutorial
110 1b we provide a generalisation of this approach to illustrate how one can calculate the
111 descriptive statistics when using a more complex design, such as when there are two
112 independent variables. In Tutorial 2a, we illustrate how one can fit Bayesian multi-level
113 regression models to RT data using the R package brms. We discuss possible link
114 functions, and plot the model-based effects of our predictors of interest. In Tutorial 2b we
115 fit Bayesian multi-level regression models to *timed* accuracy data to perform a micro-level
116 speed-accuracy tradeoff (SAT) analysis, which complements the event history analysis of
117 RT data for choice RT data. In Tutorial 3a, we illustrate how to fit the same type of
118 multilevel regression model for RT data in a frequentist framework using the R package
119 lme4. We then briefly compare and contrast these inferential frameworks when applied to
120 event history analysis. In Tutorial 3b, we illustrate how to perform the SAT analysis in a
121 frequentist framework.

122 In summary, even though event history analysis is a widely used statistical tool and
123 there already exist many excellent reviews (e.g., Blossfeld & Rohwer, 2002;
124 Box-Steffensmeier, 2004; Hosmer, Lemeshow, & May, 2011; Teachman, 1983) and tutorials
125 (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 (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)
3. measure the passage of time between time point zero and event occurrence in discrete or continuous time units.

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

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

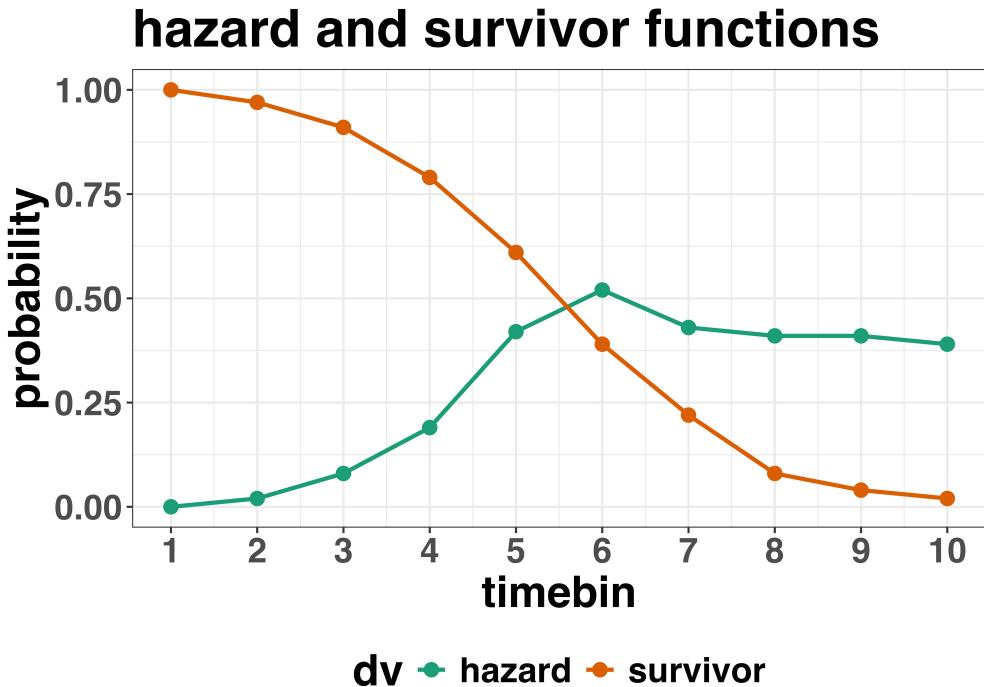


Figure 2. Discrete-time hazard and survivor functions.

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

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

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

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

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

188 2016), but simplified to make the main point clearer (Figure 3). In a standard response

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

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

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

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

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

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

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

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

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

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

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

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

201 stop and possibly interact over a particular temporal window.

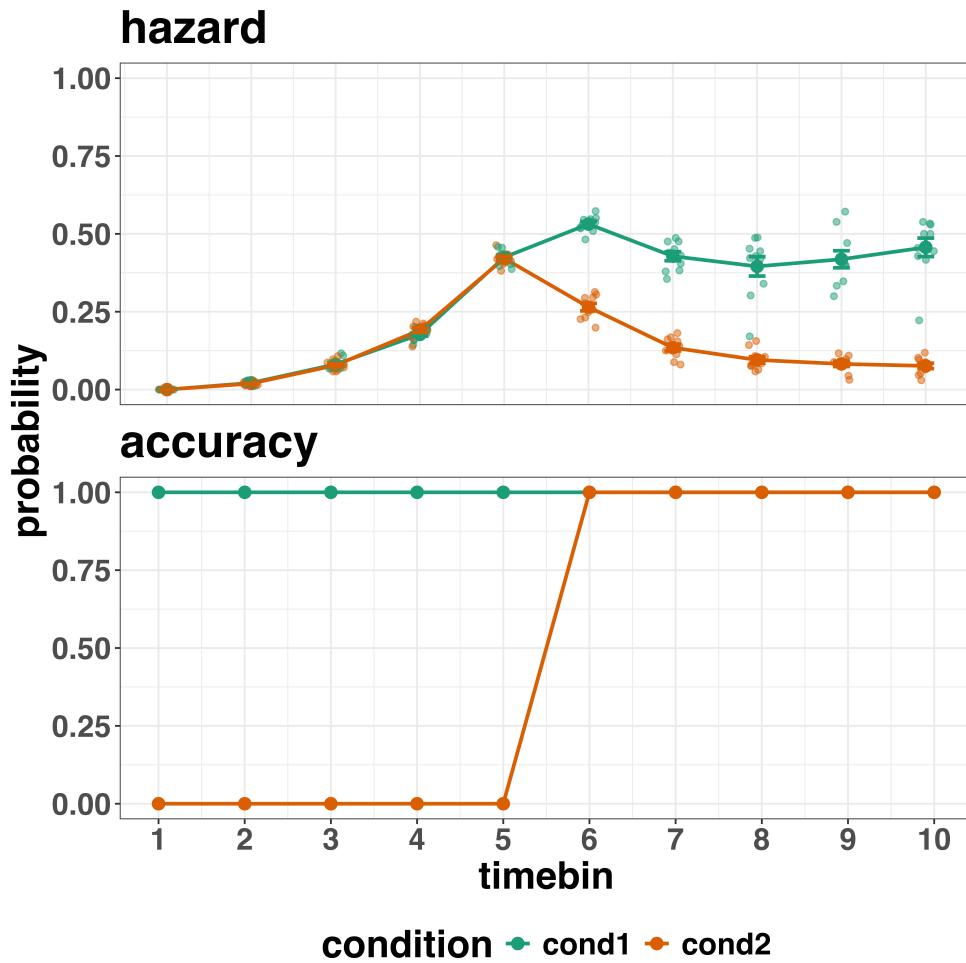


Figure 3. Discrete-time hazard and conditional accuracy functions.

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?

211 **2.2.2 Implications for designing experiments.** Performing event history
212 analyses in experimental psychology has implications for how experiments are designed.
213 Indeed, if trials are categorised as a function of when responses occur, then each timebin
214 will only include a subset of the total number of trials. For example, let's consider an
215 experiment where each participant performs 2 conditions and there are 100 trial repetitions
216 per condition. Those 100 trials must be distributed in some manner across the chosen
217 number of bins.

218 In such experimental designs, since the number of trials per condition are spread
219 across bins, it is important to have a relatively large number of trial repetitions per
220 participant and per condition. Accordingly, experimental designs using this approach
221 typically focus on factorial, within-subject designs, in which a large number of observations
222 are made on a relatively small number of participants (so-called small- N designs). This
223 approach emphasizes the precision and reproducibility of data patterns at the individual
224 participant level to increase the inferential validity of the design (Baker et al., 2021; Smith
225 & Little, 2018).

226 In contrast to the large- N design that typically average across many participants
227 without being able to scrutinize individual data patterns, small- N designs retain crucial
228 information about the data patterns of individual observers. This can be advantageous
229 whenever participants differ systematically in their strategies or in the time courses of their
230 effects, so that averaging them would lead to misleading data patterns. Note that because
231 statistical power derives both from the number of participants and from the number of
232 repeated measures per participant and condition, small- N designs can still achieve what
233 are generally considered acceptable levels of statistical power, if they have have a sufficient
234 amount of data overall (Baker et al., 2021; Smith & Little, 2018).

235 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, &

²³⁶ content of the tutorials is mainly based on Allison (2010), Singer and Willett (2003),
²³⁷ McElreath (2018), Kurz (2023a), and Kurz (2023b).

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

²³⁹ Although the focus is on EHA/SAT, we also want to briefly comment on broader
²⁴⁰ aspects of our general analytical workflow, which relate more to data science and data
²⁴¹ analysis workflows.

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

²⁴³ Descriptive, data science workflow. We perform data wrangling following tidyverse
²⁴⁴ principles and a functional programming approach (Wickham, Çetinkaya-Rundel, &
²⁴⁵ Grolemund, 2023). Functional programming basically means you don't write your own
²⁴⁶ loops but instead use functions that have been built and tested by others. [[more here, as
²⁴⁷ 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).

248 3.2 Inferential statistical approach

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

258 4. Tutorials

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

268 Our list of tutorials is as follows:

- 269 • 1a. Wrangle raw data and calculate descriptive stats for one independent variable.
- 270 • 1b. Wrangle raw data and calculate descriptive stats for two independent variables.
- 271 • 2a. Bayesian multilevel modeling for $h(t)$

- 272 • 2b. Bayesian multilevel modeling for $ca(t)$
- 273 • 3a. Frequentist multilevel modeling for $h(t)$
- 274 • 3b. Frequentist multilevel modeling for $ca(t)$

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

276 are happy with then we could include it here.

277 **4.1 Tutorial 1a: Calculating descriptive statistics using a life table**

278 **4.1.1 Data wrangling aims.** Our data wrangling procedures serve two related

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

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

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

296 **4.1.2 A real data wrangling example.** To illustrate how to quickly set up life
 297 tables for calculating the descriptive statistics (functions of discrete time), we use a
 298 published data set on masked response priming from Panis and Schmidt (2016). In their
 299 first experiment, Panis and Schmidt (2016) presented a double arrow for 94 ms that
 300 pointed left or right as the target stimulus with an onset at time point zero in each trial.
 301 Participants had to indicate the direction in which the double arrow pointed using their
 302 corresponding index finger, within 800 ms after target onset. Response time and accuracy
 303 were recorded on each trial. Prime type (blank, congruent, incongruent) and mask type
 304 were manipulated. Here we focus on the subset of trials in which no mask was presented.

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

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

310 The required column names are as follows:

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

317 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")))
```

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

```

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

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

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

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

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

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

347 The warning messages are generated because some bins have no hazard and $ca(t)$

348 estimates, and no error bars. They can thus safely be ignored. One can now inspect
349 different aspects, including the life table for a particular condition of a particular subject,
350 and a plot of the different functions for a particular participant.

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

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

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

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

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

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

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

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

375 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.

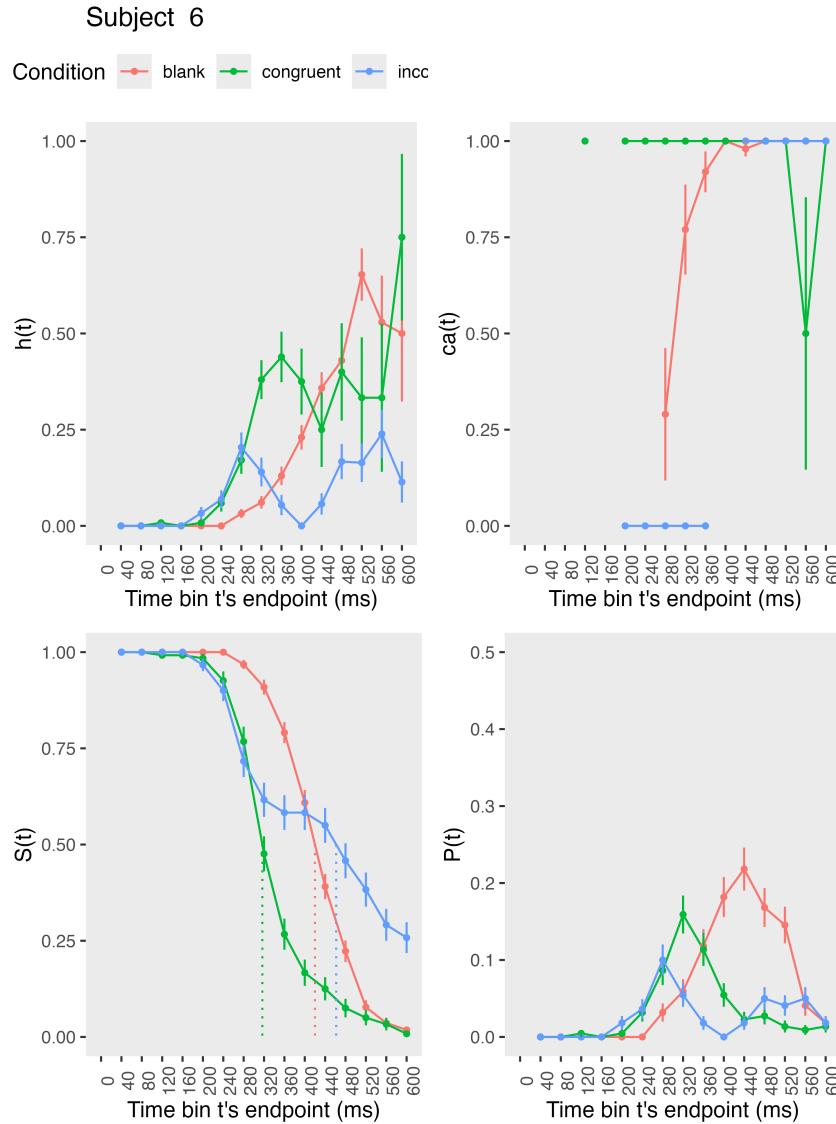


Figure 4. Estimated discrete-time hazard, survivor, probability mass, and conditional accuracy functions for participant 6, as a function of the passage of discrete waiting time.

376 However, when the waiting time has increased until 400 ms after target onset, then
 377 the conditional probability of response occurrence in the next 40 ms is estimated to be
 378 0.36, 0.25, and 0.06 for the blank, congruent, and incongruent prime conditions,
 379 respectively. And when a response does occur in bin (400,440], then the probability that it
 380 is correct is estimated to be 0.98, 1, and 1 for the blank, congruent, and incongruent prime

381 conditions, respectively.

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

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

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

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

399 **4.2 Tutorial 1b: Generalising to a more complex design**

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

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

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

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

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

410 participant 6 in Figure 5.

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

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

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

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

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

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

417 incongruent primes.

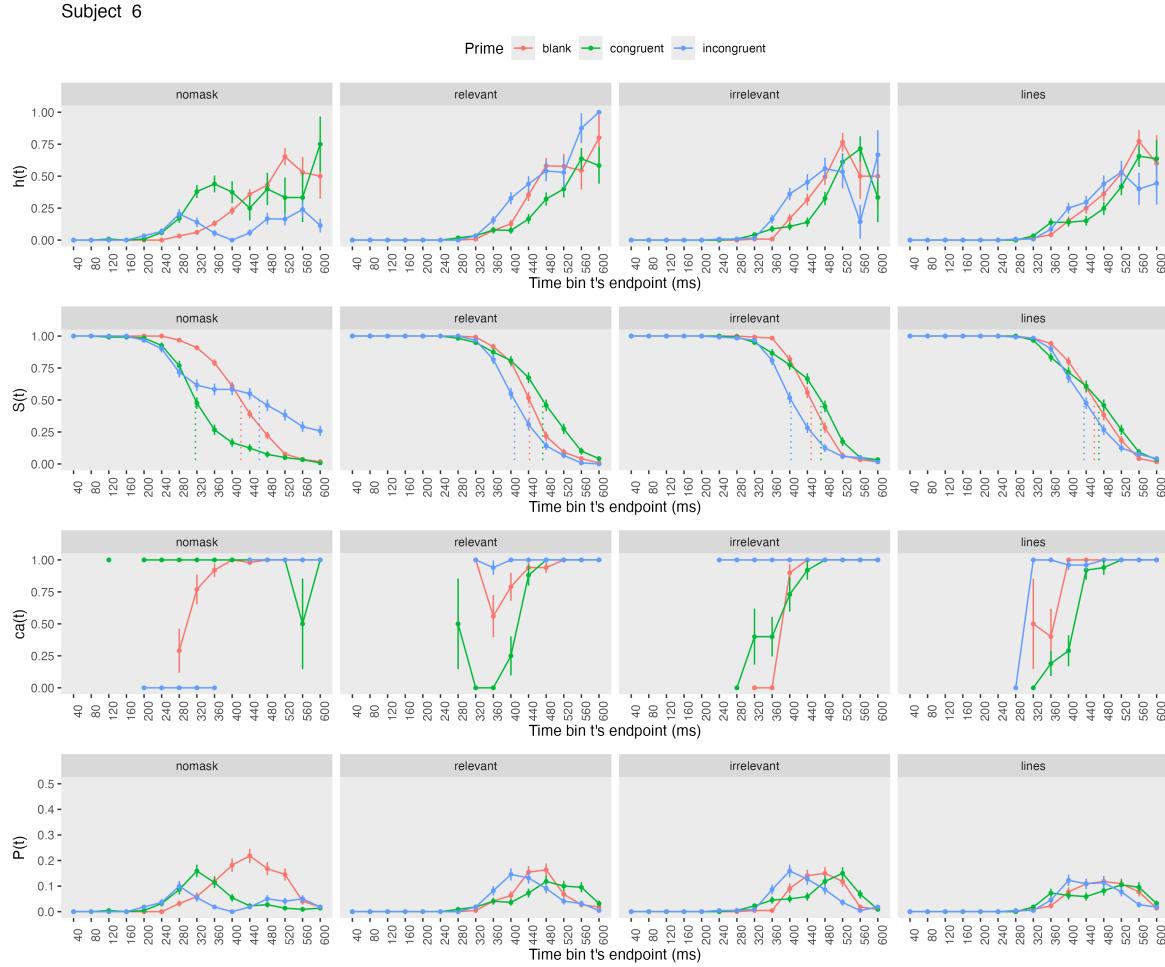


Figure 5. Sample-based discrete-time hazard, survivor, conditional accuracy, and probability mass functions for participant 6, as a function of prime type (blank, congruent, incongruent) and mask type (no mask, relevant, irrelevant, lines).

418 **4.3 Tutorial 2a: Fitting Bayesian hazard models to time-to-event data**

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

424 **4.3.1 Hazard model considerations.** There are several analytic decisions one

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

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

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

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

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

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

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

450 created in Tutorial 1a. We decided to select the analysis time window (200,600] and the
 451 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))
```

452 **4.3.2 Prior distributions.** To get the posterior distribution of each model

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

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

462 **4.3.3 Model 1: A general specification of TIME, and main effects of**
 463 **congruency and trial number.** When you do not want to make assumptions about the
 464 shape of the hazard function in the selected baseline condition, or its shape is not smooth
 465 but irregular, then you can use a general specification of TIME, i.e., fit one intercept per
 466 time bin. In this first model, we use a general specification of TIME for the selected
 467 baseline condition (blank prime), and assume that the effects of prime-target congruency
 468 and trial number are proportional and additive, and that the effect of trial number is
 469 linear. Before we fit model 1, we remove unnecessary columns from the data, and specify
 470 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")
```

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

477 **4.3.4 Model 2: A polynomial specification of TIME, and main effects of**
 478 **congruency and trial number.** When the shape of the hazard function is rather
 479 smooth, as it is for behavioral RT data, one can fit a more parsimonious model by using a

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

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

488 **4.3.5 Model 3: A polynomial specification of TIME, and relaxing the
 489 proportionality assumption.** So far, we assumed that the effect of our predictors
 490 prime-target congruency and centered trial number are the same in each time bin. However,
 491 the descriptive plots (e.g., Figure 4) suggest that the effect of prime-target congruency
 492 varies across time bins. Previous research has shown that psychological effects typically
 493 change over time (Panis, 2020; Panis, Moran, et al., 2020; Panis & Schmidt, 2022; Panis et
 494 al., 2017; Panis & Wagemans, 2009). For the third model, we thus use a third-order
 495 polynomial specification of TIME for the baseline condition (blank prime), and relax the
 496 proportionality assumption for the predictor variables prime-target congruency (variable
 497 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),
```

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

499 took about 4.5 hours.

500 4.3.6 Model 4: A polynomial specification of TIME, and relaxing all three

501 **assumptions.** Based on previous work (Panis, 2020; Panis, Moran, et al., 2020; Panis &
502 Schmidt, 2022; Panis et al., 2017; Panis & Wagemans, 2009) 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.

```
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)
```

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

507 about 8 hours.

508 4.3.7 Compare the models. We can compare the four models using the Widely

509 Applicable Information Criterion (WAIC) and Leave-One-Out (LOO) cross-validation, and

510 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)
```

```
511 ## model_M1 model_M2 model_M3 model_M4
512 ##      0      0      0      1
```

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

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

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

517 **4.3.8 Evaluate parameter estimates.** To make inferences from the parameter
 518 estimates in model M4, we summarize the draws from the posterior distributions of the
 519 effects of congruent and incongruent primes relative to the blank prime condition, in each
 520 time bin for trial numbers 500, 1000, and 1500, in terms of point and interval estimates.

521 Figure 6 shows one point (mean) and three highest posterior density interval
 522 (50/80/95%) estimates for the effects of congruent and incongruent primes relative to
 523 neutral primes, for each time bin in trial numbers 500, 1000, and 1500.

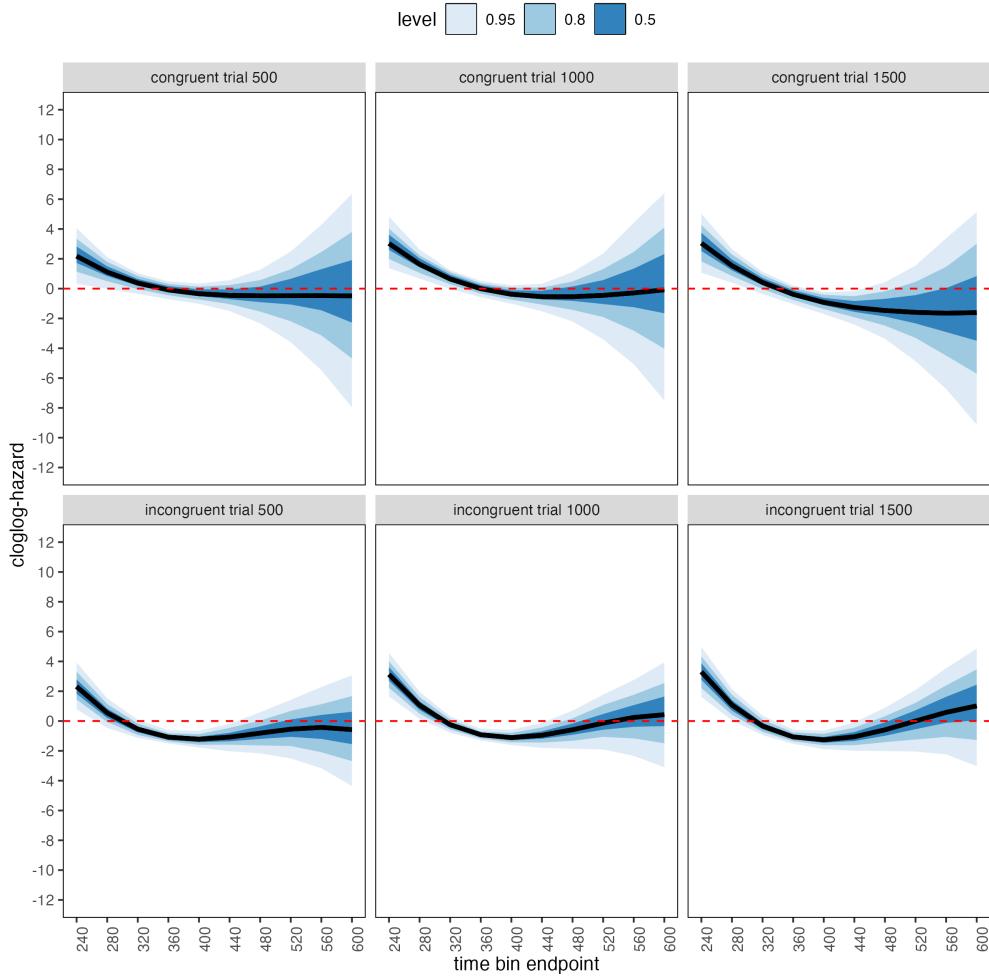


Figure 6. Means and 50/80/95% highest posterior density intervals of the draws from the posterior distributions representing the effect of congruent and incongruent primes relative to neutral primes in each bin for trial numbers 500, 1000, and 1500.

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

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

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

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

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

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

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

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

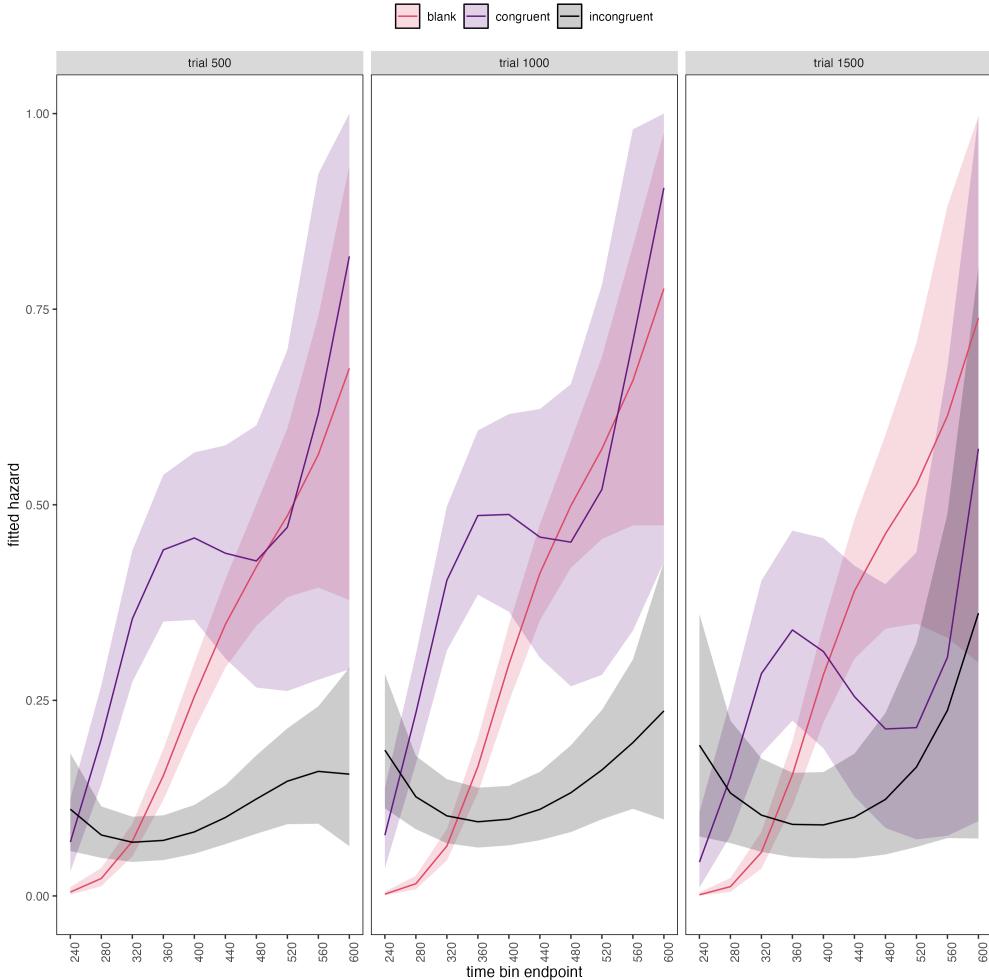


Figure 7. Model-based hazard functions for participant 6 in trial 500, 1000, and 1500.

577 **4.4 Tutorial 2b: Fitting Bayesian conditional accuracy models**

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

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

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

588 Figure 8 shows one point (mean) and three highest posterior density interval
 589 (50/80/95%) estimates for the effects of congruent and incongruent primes relative to
 590 neutral primes on logit-ca, for each time bin in trial numbers 500, 1000, and 1500.

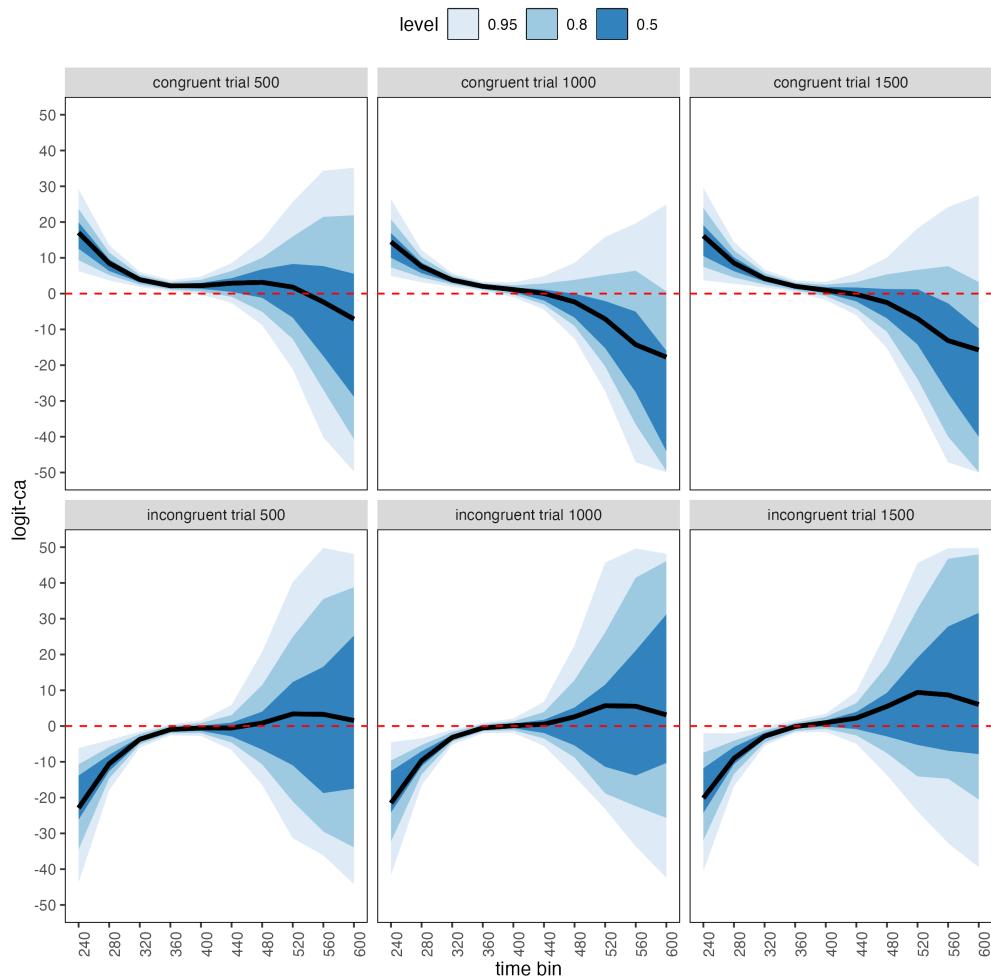


Figure 8. Means and 50/80/95% highest posterior density intervals of the draws from the posterior distributions representing the effect of congruent and incongruent primes relative to neutral primes in each bin for trial numbers 500, 1000, and 1500.

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

592 effects of congruent and incongruent primes relative to the blank prime condition in trials
 593 500, 1000, and 1500, in terms of a point estimate (the mean) and the upper and lower
 594 bounds of the 95% highest posterior density interval. Also, by exponentiating the mean we
 595 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.

596

597 Based on Figure 8 and Table 5, we see that throughout the experiment (trials 500,
598 1000, and 1500), congruent primes have a positive effect on logit-ca(t) in time bins
599 (200,240], (240,280], (280,320], and (320,360], relative to the logit-ca(t) estimates in the
600 baseline condition (blank prime; red striped lines in Figure 8). For example, the odds ratio
601 for congruent primes in bin (320,360] in trial 500 shows that the odds of a correct response
602 are estimated to be 8.89 times higher than the odds of a correct response when there is no
603 prime. Incongruent primes have a negative effect on logit-ca(t) in time bins (200,240],
604 (240,280], and (280,320] throughout the experiment, relative to the logit-ca(t) estimates in
605 the baseline condition (no prime; red striped lines).

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

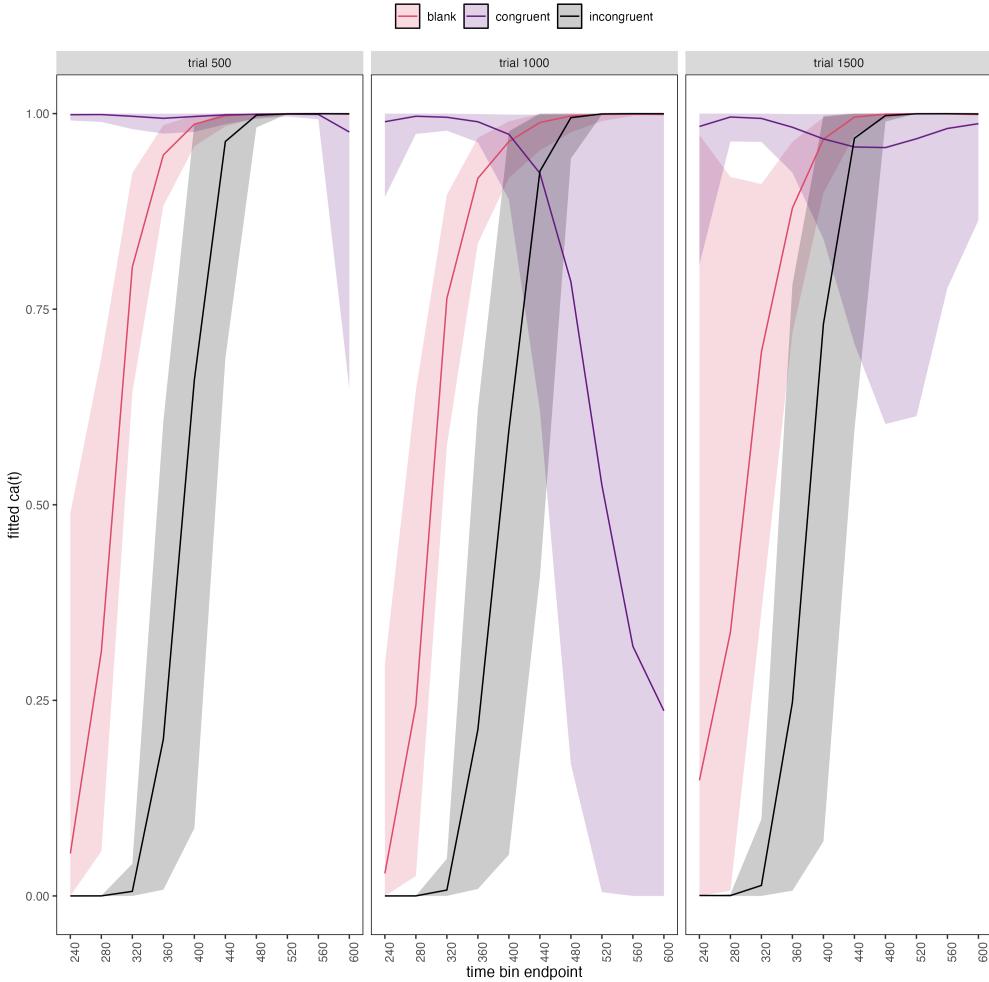


Figure 9. Model-based $ca(t)$ functions for participant 6 in trial 500, 1000, and 1500.

608 4.5 Tutorial 3a: Fitting Frequentist hazard models

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

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

617 of model M3_f from `glmer()`.

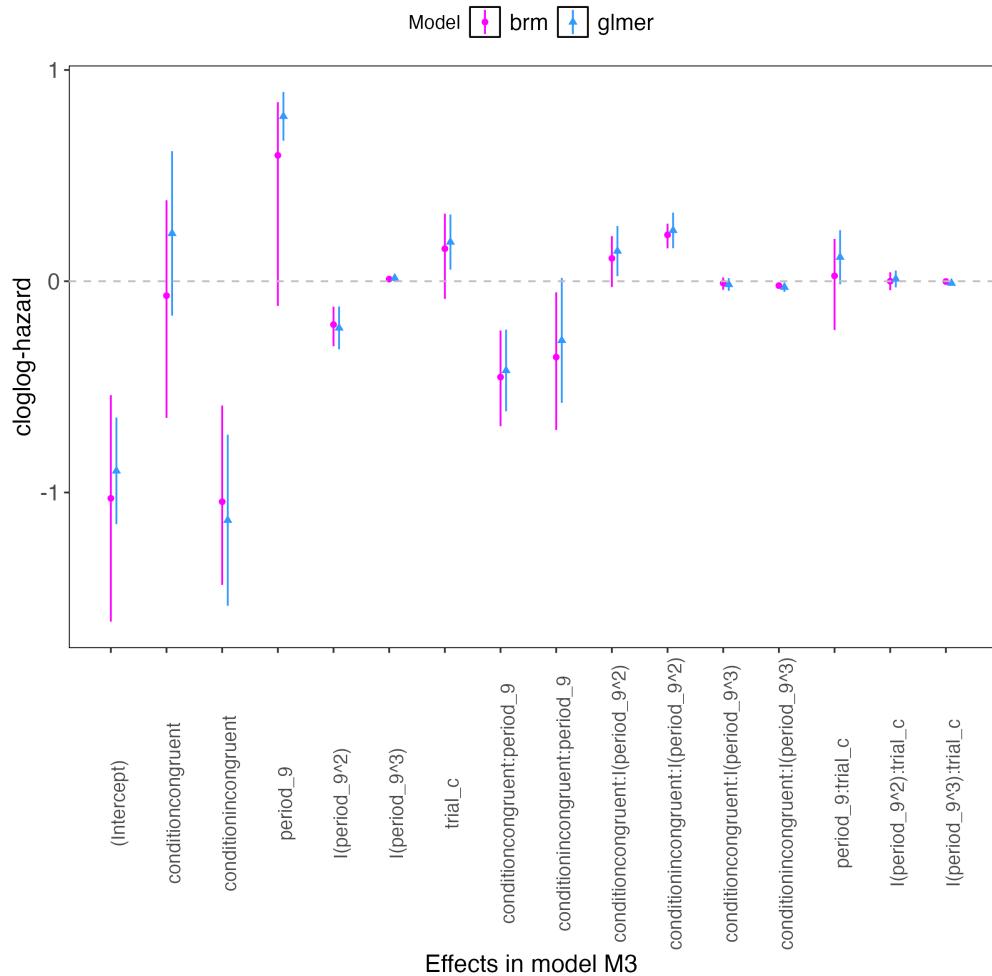


Figure 10. Parameter estimates for model M3 from `brm()` – means and 95% quantile intervals – and model M3_f from `glmer()` – maximum likelihood estimates and 95% confidence intervals.

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

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

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

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

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

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

626 **4.6 Tutorial 3b: Fitting Frequentist conditional accuracy models**

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

633 **5. Discussion**

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

648 5.1 Advantages of hazard analysis

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

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

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

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

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

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

691 5.2 Model complexity versus interpretability

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

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

706 **5.3 Individual differences**

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

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

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

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

733 5.4 Limitations

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

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

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

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

759 **5.5 Extensions**

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

775

6. Conclusions

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

787

References

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

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

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

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

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

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

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

953

Supplementary material

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

956 The shape of a distribution of waiting times can be described in multiple ways (Luce,
 957 1991). After dividing time in discrete, contiguous time bins indexed by t , let RT be a
 958 discrete random variable denoting the rank of the time bin in which a particular person's
 959 response occurs in a particular experimental condition. Because waiting times can only
 960 increase, discrete-time EHA focuses on the discrete-time hazard function

$$961 \quad h(t) = P(RT = t | RT \geq t) \quad (1)$$

962 and the discrete-time survivor function

$$963 \quad S(t) = P(RT > t) = [1-h(t)].[1-h(t-1)].[1-h(t-2)] \dots [1-h(1)] \quad (2)$$

964 and not on the probability mass function

$$965 \quad P(t) = P(RT = t) = h(t).S(t-1) \quad (3)$$

966 nor the cumulative distribution function

$$967 \quad F(t) = P(RT \leq t) = 1-S(t) \quad (4)$$

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

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

977 discrete-time conditional accuracy function

978 $ca(t) = P(\text{correct} \mid RT = t)$ (5)

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

982 The survivor function provides a context for the hazard function, as $S(t-1) = P(RT >$
 983 $t-1) = P(RT \geq t)$ tells you on how many percent of the trials the estimate $h(t) = P(RT =$
 984 $t \mid RT \geq t)$ is based. The probability mass function provides a context for the conditional
 985 accuracy function, as $P(t) = P(RT = t)$ tells you on how many percent of the trials the
 986 estimate $ca(t) = P(\text{correct} \mid RT = t)$ is based.

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

989 Continuous-time EHA does not focus on the cumulative distribution function $F(t) = P(RT$
 990 $\leq t)$ and its derivative, the probability density function $f(t) = F(t)'$, but on the survivor
 991 function $S(t) = P(RT > t)$ and the hazard rate function $\lambda(t) = f(t)/S(t)$. The hazard rate
 992 function gives you the instantaneous *rate* of event occurrence at time point t , given that
 993 the event has not occurred yet.

994 **B. Custom functions for descriptive discrete-time hazard analysis**

995 We defined 13 custom functions that we list here.

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

- 1000 • setup_lt_2IV(ptb) : set up a life table for each combination of levels of 2
1001 independent variables
- 1002 • calc_ca(df) : estimate the conditinal accuracies when there is 1 independent variable
- 1003 • calc_ca_2IV(df) : estimate the conditional accuraiies when there are 2 independent
1004 variables
- 1005 • join_lt_ca(df1,df2) : add the ca(t) estimates to the life tables (1 independent
1006 variable)
- 1007 • join_lt_ca_2IV(df1,df2) : add the ca(t) estimates to the life tables (2 independent
1008 variables)
- 1009 • extract_median(df) : estimate quantiles $S(t)._{50}$ (1 independent variable)
- 1010 • extract_median_2IV(df) : estimate quantiles $S(t)._{50}$ (2 independent variables)
- 1011 • plot_eha(df,subj,haz_yaxis) : create plots of the discrete-time functions (1
1012 independent variable)
- 1013 • plot_eha_2IV(df,subj,haz_yaxis) : create plots of the discrete-time functions (2
1014 independent variables)
- 1015 • plot_eha_agg(df,subj,haz_yaxis) : create 1 plot for data aggregated across
1016 participants (1 independent variable)

1017 When you want to analyse simple RT data from a detection experiment with one
1018 independent variable, the functions calc_ca() and join_lt_ca() should not be used, and
1019 the code to plot the conditional accuracy functions should be removed from the function
1020 plot_eha(). When you want to analyse simple RT data from a detection experiment with
1021 two independent variables, the functions calc_ca_2IV() and join_lt_ca_2IV() should not
1022 be used, and the code to plot the conditional accuracy functions should be removed from
1023 the function plot_eha_2IV().

1024 **C. Link functions**

1025 Popular link functions include the logit link and the complementary log-log link, as
1026 shown in Figure 11.

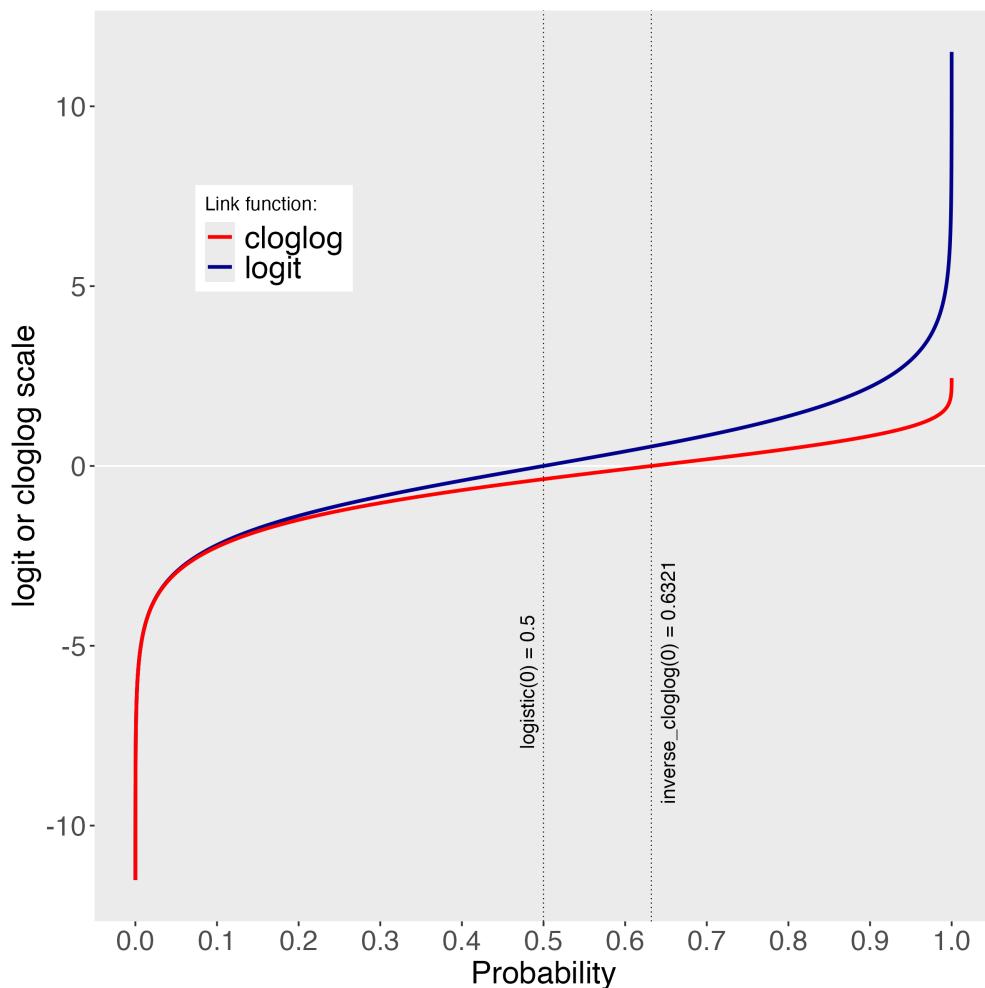


Figure 11. The logit and cloglog link functions.

1027 **D. Regression equations**

1028 An example (single-level) discrete-time hazard model with three predictors (TIME,
1029 X₁, X₂), the cloglog link function, and a third-order polynomial specification for TIME can
1030 be written as follows:

$$\begin{aligned} 1031 \quad & \text{cloglog}[h(t)] = \ln(-\ln[1-h(t)]) = [\beta_0 \text{ONE} + \beta_1(\text{TIME}-9) + \beta_2(\text{TIME}-9)^2] + [\beta_3 X_1 + \beta_4 X_2 \\ 1032 \quad & + \beta_5 X_2(\text{TIME}-9)] \end{aligned} \quad (6)$$

1033 The main predictor variable TIME is the time bin index t that is centered on value 9
 1034 in this example. The first set of terms within brackets, the parameters β_0 to β_2 multiplied
 1035 by their polynomial specifications of (centered) time, represents the shape of the baseline
 1036 cloglog-hazard function (i.e., when all predictors X_i take on a value of zero). The second
 1037 set of terms (the beta parameters β_3 to β_5) represents the vertical shift in the baseline
 1038 cloglog-hazard for a 1 unit increase in the respective predictor variable. Predictors can be
 1039 discrete, continuous, and time-varying or time-invariant. For example, the effect of a 1 unit
 1040 increase in X_1 is to vertically shift the whole baseline cloglog-hazard function by β_1
 1041 cloglog-hazard units. However, if the predictor interacts linearly with TIME (see X_2 in the
 1042 example), then the effect of a 1 unit increase in X_2 is to vertically shift the predicted
 1043 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
 1044 predictor, its β parameter is exponentiated, resulting in a hazard ratio (due to the use of
 1045 the cloglog link). When using the logit link, exponentiating a β parameter results in an
 1046 odds ratio.
 1047 odds ratio.

1048 An example (single-level) discrete-time hazard model with a general specification for
 1049 TIME (separate intercepts for each of six bins, where D1 to D6 are binary variables
 1050 identifying each bin) and a single predictor (X_1) can be written as follows:

$$1051 \quad \text{cloglog}[h(t)] = [\beta_0 D1 + \beta_1 D2 + \beta_2 D3 + \beta_3 D4 + \beta_4 D5 + \beta_5 D6] + [\beta_6 X_1] \quad (7)$$

1052 E. Prior distributions

1053 To gain a sense of what prior *logit* values would approximate a uniform distribution
 1054 on the probability scale, Kurz (2023a) simulated a large number of draws from the
 1055 Uniform(0,1) distribution, converted those draws to the log-odds metric, and fitted a

1056 Student's t distribution. Row C in Figure 12 shows that using a t-distribution with 7.61
 1057 degrees of freedom and a scale parameter of 1.57 as a prior on the logit scale, approximates
 1058 a uniform distribution on the probability scale. According to Kurz (2023a), such a prior
 1059 might be a good prior for the intercept(s) in a logit-hazard model, while the $N(0,1)$ prior in
 1060 row D might be a good prior for the non-intercept parameters in a logit-hazard model, as it
 1061 gently regularizes p towards .5 (i.e., a zero effect on the logit scale).

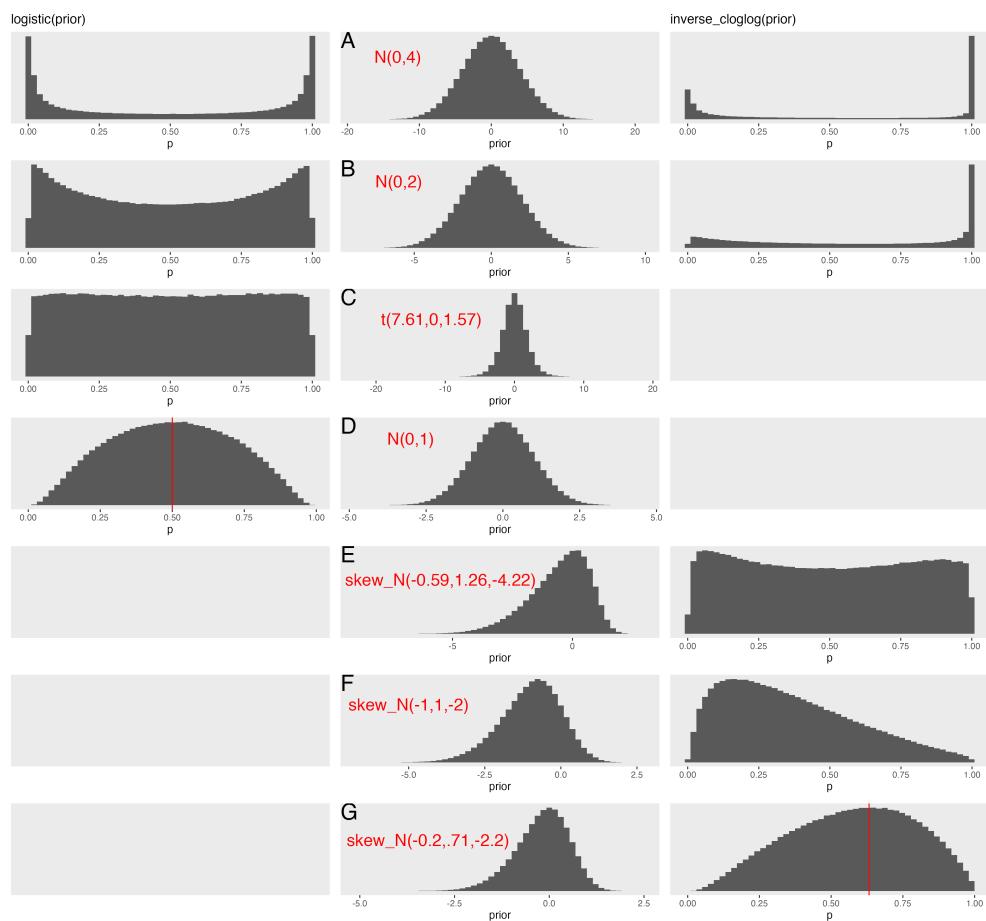


Figure 12. Prior distributions on the logit and/or cloglog scales (middle column), and their implications on the probability scale after applying the inverse-logit (or logistic) transformation (left column), and the inverse-cloglog transformation (right column).

1062 To gain a sense of what prior *cloglog* values would approximate a uniform distribution
 1063 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).