

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

3 Sven Panis<sup>1</sup> & Richard Ramsey<sup>1</sup>

4 <sup>1</sup> 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, Gloriastrasse 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

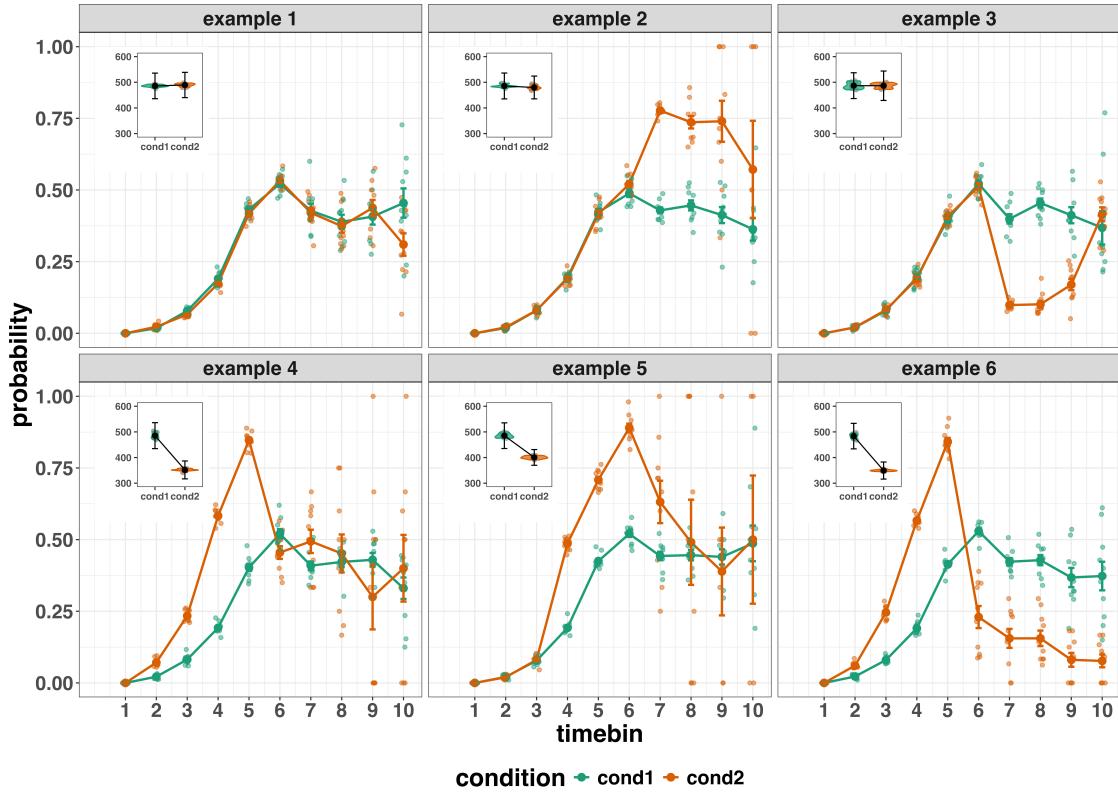
40 Event History Analyses for psychological time-to-event data: A tutorial in R with examples  
41 in Bayesian and frequentist workflows

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 of 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). Such  
54 information is useful not only for interpretation of the effects, but also for cognitive  
55 psychophysiology and computational model selection (Panis, Schmidt, Wolkersdorfer, &  
56 Schmidt, 2020).

57 As a simple illustration, Figure 1 shows the results of several simulated RT datasets,  
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 distribution differs (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 distribution differs 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 distribution would lead to a different conclusion in every case.



*Figure 1.* Means versus distributional shapes for six different simulated dataset examples. 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 timecourse of  
 67 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 waiting 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 (William Matthew Makeham, 1860). The same logic has been applied to  
84 psychological time-to-event data, as previously demonstrated (Panis et al., 2020). Here, in  
85 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 .html file, so that  
106 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 models 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 reflects 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.

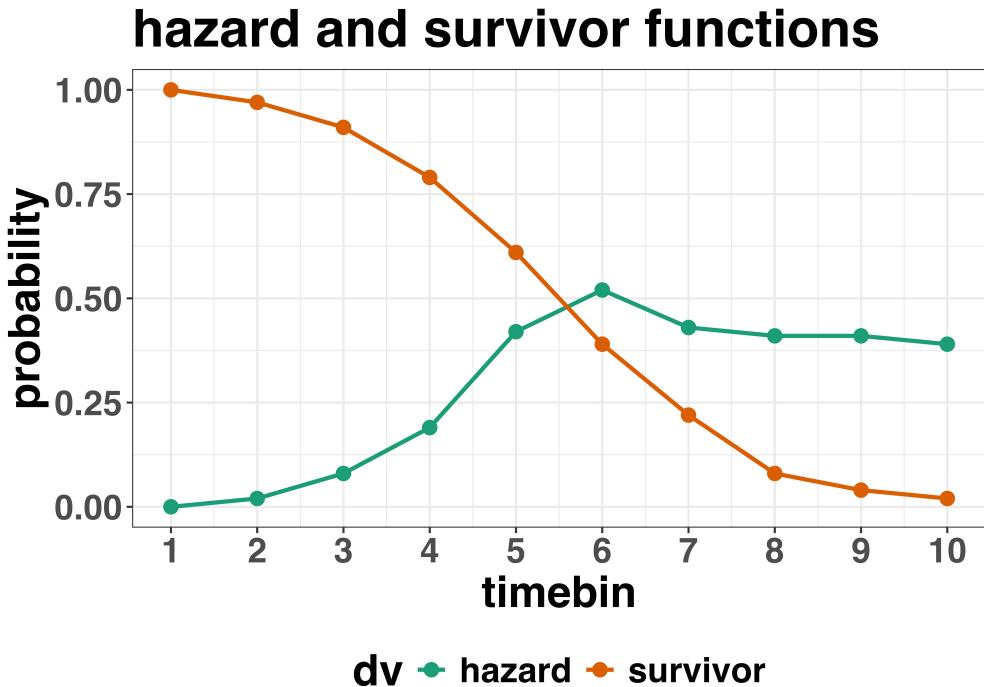


Figure 2. Hazard and survivor functions

<sup>172</sup> **2.2 Event history analysis in the context of experimental psychology**

<sup>173</sup> **2.2.1 A worked example.** In the context of experimental psychology, it is  
<sup>174</sup> common for participants to be presented with either a 1-button detection task or a  
<sup>175</sup> 2-button discrimination task, i.e., a task that has a right and a wrong answer. For  
<sup>176</sup> example, a task may involve choosing between two response options with only one of them  
<sup>177</sup> being correct. For such two-choice RT data, the discrete-time EHA can be extended with a  
<sup>178</sup> discrete-time SAT analysis. Specifically, the hazard function of event occurrence can be  
<sup>179</sup> extended with the discrete-time conditional accuracy function (see equation 5 in part A of  
<sup>180</sup> the supplementary material), which gives you the probability that a response is correct  
<sup>181</sup> given that it is emitted in time bin  $t$  (Allison, 2010; Kantowitz & Pachella, 2021;  
<sup>182</sup> Wickelgren, 1977). We refer to this extended analysis for choice RT data as EHA/SAT. As  
<sup>183</sup> explained in part A of the supplementary material, the probability mass function provides  
<sup>184</sup> 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 example that is inspired by real data (Panis et al., 2016), but  
188 simplified to make the main point clearer (Figure 3). In a standard response priming  
189 paradigm, there is a prime stimulus (e.g., an arrow pointing left or right) followed by a  
190 target stimulus (another arrow pointing left or right). The prime can then be congruent or  
191 incongruent with the target. Figure 3 shows that the early upswing in hazard is equal for  
192 both prime conditions, and that early responses are always correct in the congruent  
193 condition and always incorrect in the incongruent condition. These results show that for  
194 short waiting times (< bin 6), responses always follow the prime (and not the target, as  
195 instructed). And then for longer waiting times, response hazard is lower in incongruent  
196 compared to congruent trials, and all responses emitted in these later bins are correct. This  
197 is interesting because mean-average RT would only represent the overall ability of cognition  
198 to overcome interference, on average, across trials. And such a conclusion is not supported  
199 when the effects are explored over a timeline. Instead, the psychological conclusion is much  
200 more nuanced and suggests that multiple states start, stop and possibly interact over a  
201 particular temporal window.

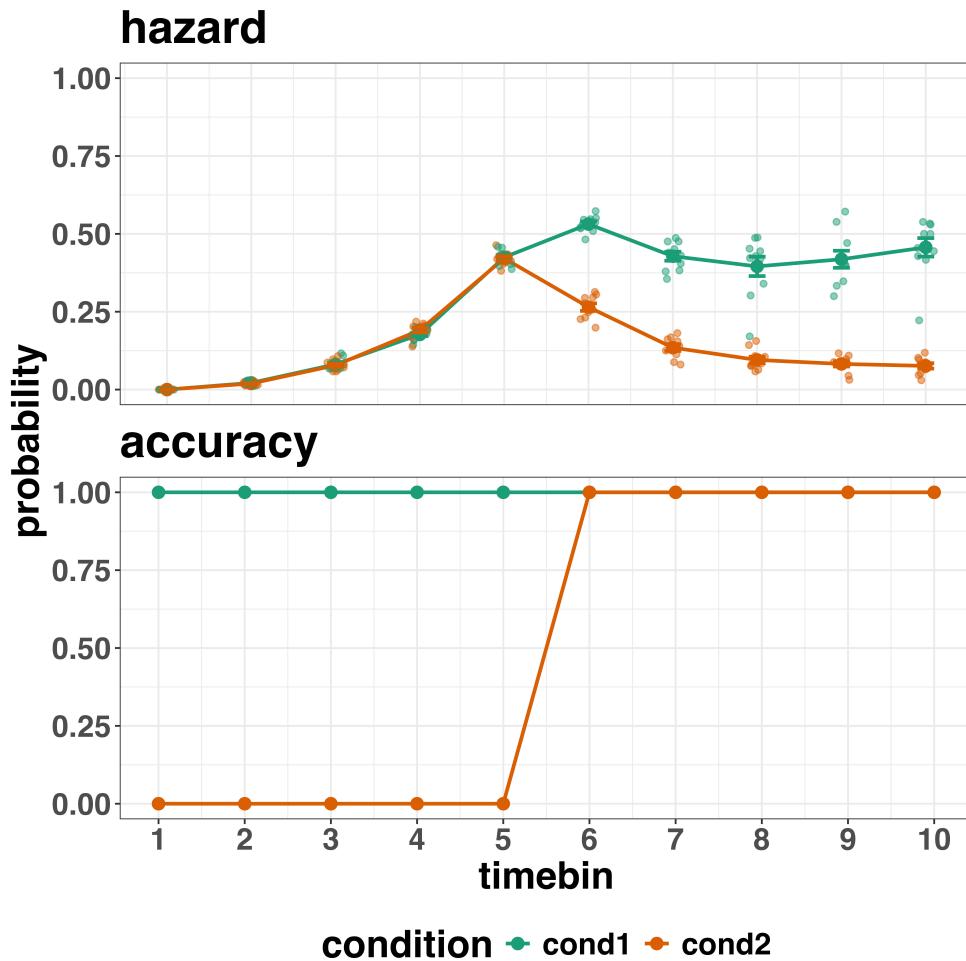


Figure 3. Hazard and conditional accuracy

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)<sup>1</sup> for all reported analyses. The

---

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

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

<sup>238</sup> **3. An overview of the general analytical workflow**

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

<sup>242</sup> **3.1 Data science workflow and descriptive statistics**

<sup>243</sup> Descriptive, data science workflow. Data wrangling via tidyverse principles and a  
<sup>244</sup> functional programming approach (cite R4DS textbook here). Functional programming  
<sup>245</sup> basically means you don't write your own loops but instead use functions that have been  
<sup>246</sup> built and tested by others. [[more here, as 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), *tidyrr* (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).

### 247 3.2 Inferential statistical approach

248 Our lab adopts a estimation approach to multi-level regression (Kruschke & Liddel,  
249 2018; Winter, 2019), which is heavily influenced by Bayesian approach as suggested by  
250 Richard McElreath (McElreath, 2020; Kurz, 202?). We also use a “keep it maximal”  
251 approach to specifying varying (or random) effects (Barr et al., 2013). This means that  
252 wherever possible we include varying intercepts and slopes per participant To make  
253 inferences, we use two main approaches. We compare models of different complexity, using  
254 information criteria, such as WAIC or LOO, to evaluate out-of-sample predictive accuracy.  
255 We also take the most complex model and evaluate key parameters of interest using point  
256 and interval estimates.

## 257 4. Tutorials

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

267 Our list of tutorials is as follows:

- 268 • 1a. Wrangle raw data and calculate descriptive stats for one independent variable.
- 269 • 1b. Wrangle raw data and calculate descriptive stats for two independent variables.
- 270 • 2a. Bayesian multilevel modeling for  $h(t)$

- 271     • 2b. Bayesian multilevel modeling for  $ca(t)$
- 272     • 3a. Frequentist multilevel modeling for  $h(t)$
- 273     • 3b. Frequentist multilevel modeling for  $ca(t)$

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

275       are happy with then we could include it here.

276   **4.1 Tutorial 1a: Calculating descriptive statistics using a life table**

277       **4.1.1 Data wrangling aims.** Our data wrangling procedures serve two related

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

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

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

295       **4.1.2 A real data wrangling example.** To illustrate how to quickly set up life  
 296 tables for calculating the descriptive statistics (functions of discrete time), we use a  
 297 published data set on masked response priming from Panis and Schmidt (2016). In their  
 298 first experiment, Panis and Schmidt (2016) presented a double arrow for 94 ms that  
 299 pointed left or right as the target stimulus with an onset at time point zero in each trial.  
 300 Participants had to indicate the direction in which the double arrow pointed using their  
 301 corresponding index finger, within 800 ms after target onset. Response time and accuracy  
 302 were recorded on each trial. Prime type (blank, congruent, incongruent) and mask type  
 303 were manipulated. Here we focus on the subset of trials in which no mask was presented.

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

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

309 The required column names are as follows:

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

316 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")))
```

317 Next, we can set up the life tables and plots of the discrete-time functions  $h(t)$ ,  $S(t)$ ,  
 318  $ca(t)$ , and  $P(t)$  – see part A of the supplementary material for their definitions. To do so  
 319 using a functional programming approach, one has to nest the data within participants  
 320 using the group\_nest() function, and supply a user-defined censoring time and bin width  
 321 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 dataset
mutate(lifetable = map(ptb_data, setup_lt)) %>%    # create life tables without ca(t)
mutate(condacc = map(censored, calc_ca)) %>%      # calculate ca(t)
mutate(lifetable_ca = map2(lifetable, condacc, join_lt_ca)) %>%    # create life tables with ca(t)
mutate(plot = map2(.x = lifetable_ca, .y = pid, plot_eha,1)) # create plots

```

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

374 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
0.00	220.00	NA	NA	NA	1.00	0.00	NA	NA
40.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
80.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
120.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
160.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
200.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
240.00	220.00	0.00	0.00	0.00	1.00	0.00	NA	NA
280.00	220.00	7.00	0.03	0.01	0.97	0.01	0.29	0.17
320.00	213.00	13.00	0.06	0.02	0.91	0.02	0.77	0.12
360.00	200.00	26.00	0.13	0.02	0.79	0.03	0.92	0.05
400.00	174.00	40.00	0.23	0.03	0.61	0.03	1.00	0.00
440.00	134.00	48.00	0.36	0.04	0.39	0.03	0.98	0.02
480.00	86.00	37.00	0.43	0.05	0.22	0.03	1.00	0.00
520.00	49.00	32.00	0.65	0.07	0.08	0.02	1.00	0.00
560.00	17.00	9.00	0.53	0.12	0.04	0.01	1.00	0.00
600.00	8.00	4.00	0.50	0.18	0.02	0.01	1.00	0.00

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

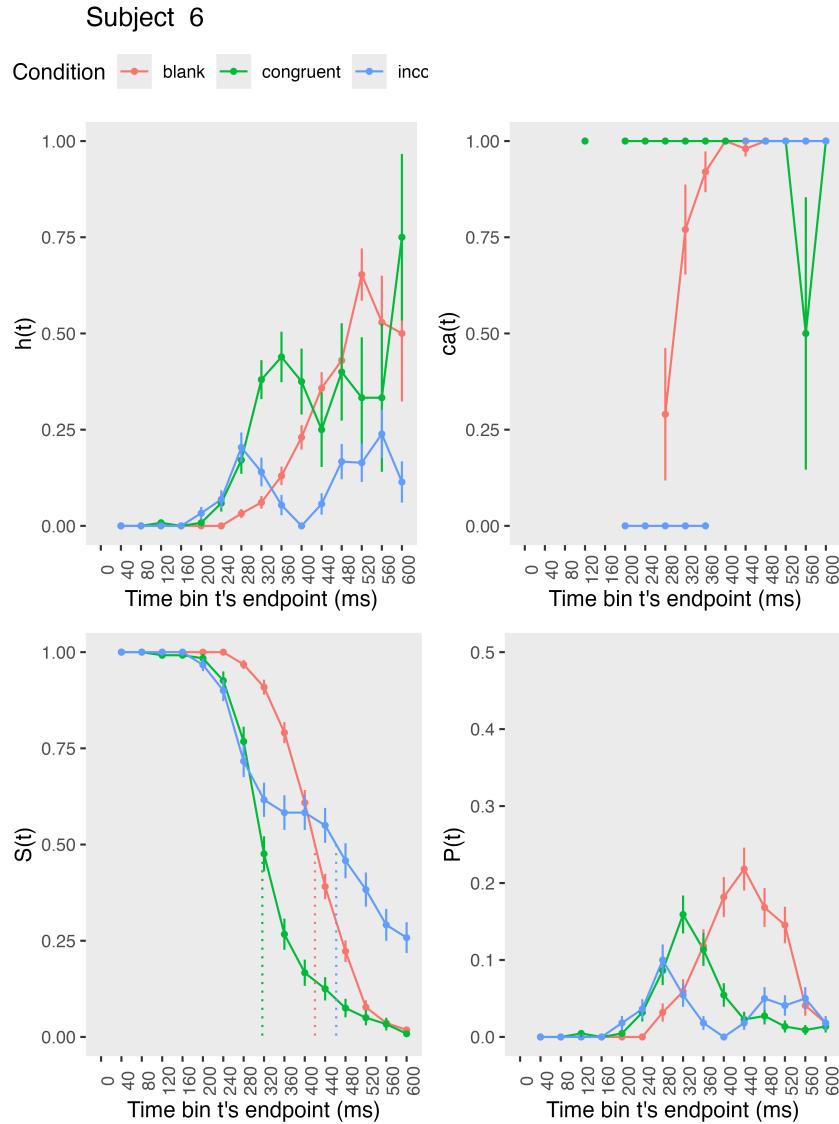


Figure 4. Estimated discrete-time hazard, survivor, and conditional accuracy functions for participant 6, as a function of the passage of discrete waiting time.

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

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

377 0.36, 0.25, and 0.06 for the blank, congruent, and incongruent prime conditions,

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

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

380 conditions, respectively.

381 When participants show qualitatively the same distributional patterns, one might  
382 consider to aggregate their data and make one plot (see Tutorial\_1a.Rmd).

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

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

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

## 398 **4.2 Tutorial 1b: Generalising to a more complex design**

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

405 To this end, Tutorial 1b illustrates how to calculate and plot the descriptive statistics  
406 for the full data set of Experiment 1 of Panis and Schmidt (2016), which includes two  
407 independent variables: mask type and prime type. As we use the same functional  
408 programming approach as in Tutorial 1a, we simply present the sample-based functions for  
409 participant 6 in Figure 5.

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

411 effect in the hazard and  $ca(t)$  functions, as congruent primes temporarily generate higher  
412 values for hazard and conditional accuracy compared to incongruent primes. However,

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

414 compatibility effect in the hazard and conditional accuracy functions, as congruent primes  
415 temporarily generate *lower* values for hazard and conditional accuracy compared to

416 incongruent primes.

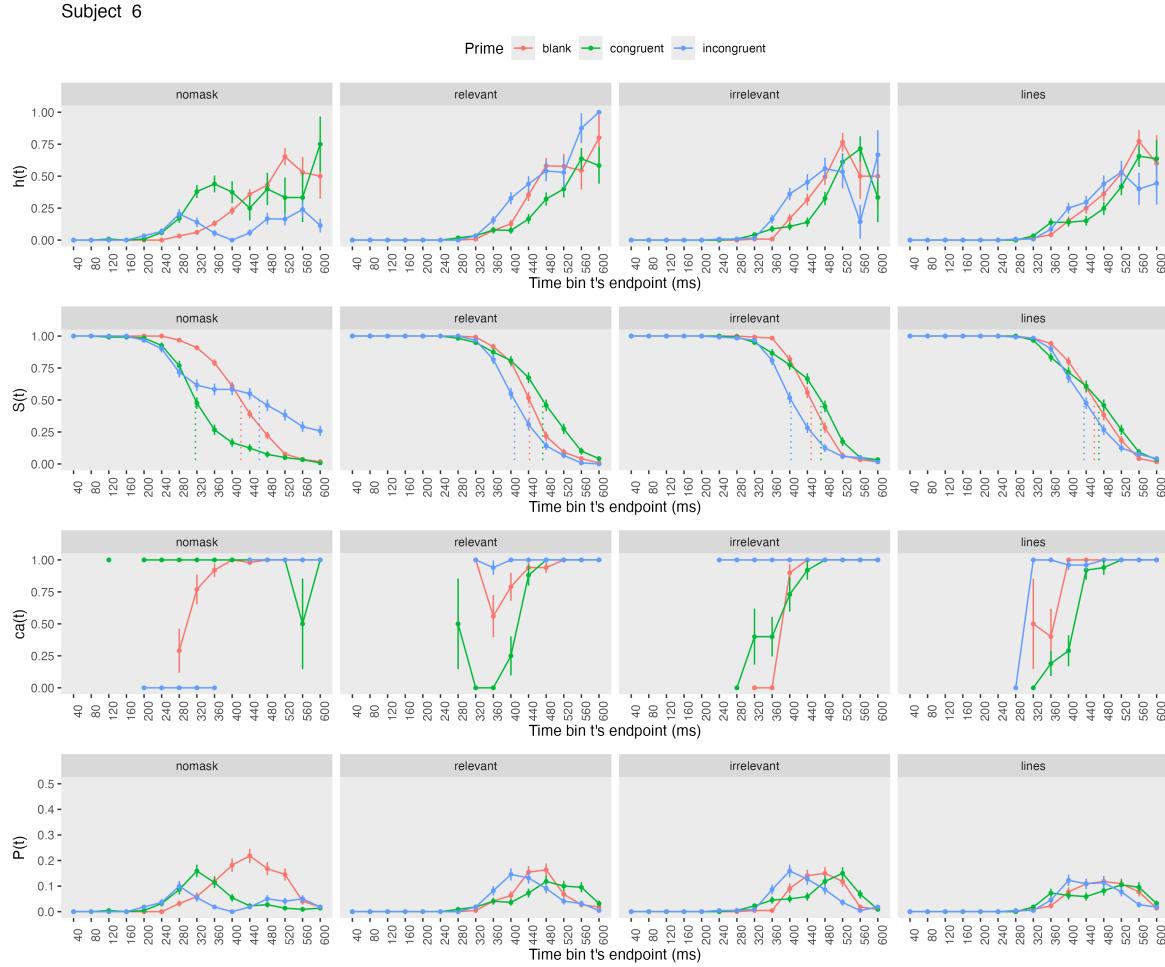


Figure 5. Sample-based discrete-time hazard, survivor, conditional accuracy, and probability mass functions for participant 6, as a function of prime type (blank, congruent, incongruent) and mask type (no mask, relevant, irrelevant, lines).

417 **4.3 Tutorial 2a: Fitting Bayesian hazard models to time-to-event data**

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

423       **4.3.1 Hazard model considerations.** There are several analytic decisions one

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

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

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

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

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

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

447       mixed models) that differ in complexity to the person-trial-bin oriented data set that we  
448       created in Tutorial 1a. We select the analysis range (200,600] and the cloglog link. The

449 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 range: (200,600] with 10 bins (time bin ranks 6 to 15)
ptb_data <- ptb_data %>% filter(period > 5)

# create factor condition, with "blank" as the reference level
ptb_data <- ptb_data %>% mutate(condition = factor(condition, labels = c("blank", "congruent", "incongruent")))

# center TIME (period) on bin 9, and trial on trial 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))
```

450       **4.3.2 Prior distributions.** To get the posterior distribution of each model

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

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

460       **4.3.3 Model 1: A general specification of TIME, and main effects of**

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

```
plan(multicore)

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

469       After selecting the binomial family and the cloglog link, the model formula is

470       specified. The fixed effects include 9 dummy variables, the explicit Intercept variable  
 471       (which represents bin 9 in this example), and the main effects of priming condition and  
 472       centered trial number. Each of these effects is allowed to vary across individuals (variable  
 473       pid). Estimating model M1 took about 70 minutes on a MacBook Pro (Sonoma 14.6.1 OS,  
 474       18GB Memory, M3 Pro Chip).

475       **4.3.4 Model 2: A polynomial specification of TIME, and main effects of**

476       **congruency and trial number.** When the shape of the hazard function is rather

smooth, as it is for behavioral RT data, one can fit a more parsimonious model by using a polynomial specification of TIME. For our second example model, we thus use a third-order polynomial specification of TIME for the baseline condition (blank prime), and again assume that the effects of prime-target congruency and centered trial number are proportional and additive, and that the effect of trial number is linear. The 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),
```

Because TIME is centered on bin 9, and trial number on trial 1000, the Intercept represents the cloglog-hazard in bin 9 for the blank prime condition in trial 1000. Estimating model M2 took about 144 minutes.

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

```

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

497 took about 268 minutes.

#### 498 4.3.6 Model 4: A polynomial specification of TIME, and relaxing all three

499 assumptions. Based on previous work (Panis, 2020; Panis, Moran, et al., 2020; Panis &  
500 Schmidt, 2022; Panis et al., 2017; Panis & Wagemans, 2009) we expect nonlinear effects of  
501 trial number. We thus relax all three assumptions in model 4. We add a squared term for  
502 the continuous predictor centered trial number –  $I(trial\_c^2)$  – and include interaction

503 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)

```

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

505 about 8 hours.

#### 506 4.3.7 Compare the models. We can compare the four models using the Widely

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

508 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)
```

```
509 ## model_M1 model_M2 model_M3 model_M4
510 ##      0      0      0      1
```

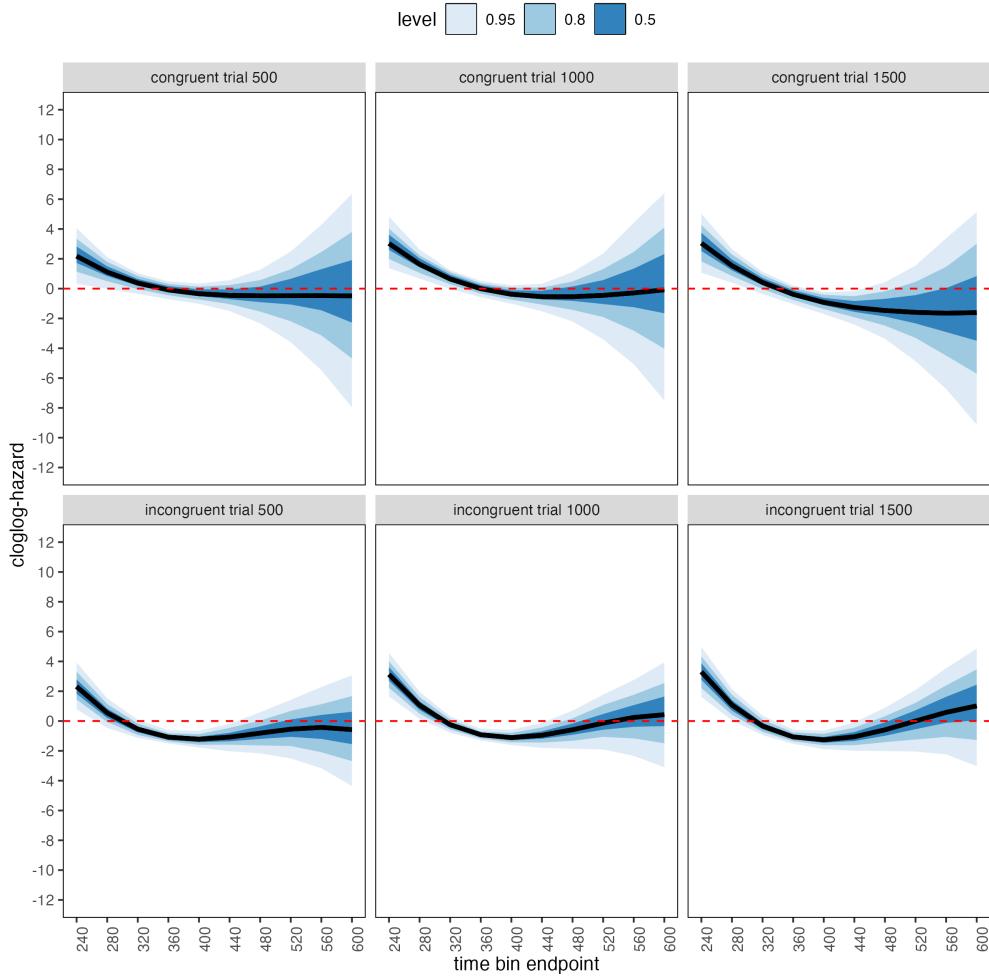
```
model_weights(model_M1, model_M2, model_M3, model_M4, weights = "waic") %>% round(digits = 3)
```

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

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

515 **4.3.8 Evaluate parameter estimates.** To make inferences from the parameter  
 516 estimates in model M4, we summarize the draws from the posterior distributions of the  
 517 effects of congruent and incongruent primes relative to the blank prime condition, in each  
 518 time bin for trial numbers 500, 1000, and 1500, in terms of point and interval estimates.

519 Figure 6 shows one point (mean) and three highest posterior density interval  
 520 (50/80/95%) estimates for the effects of congruent and incongruent primes relative to  
 521 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.

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

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

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

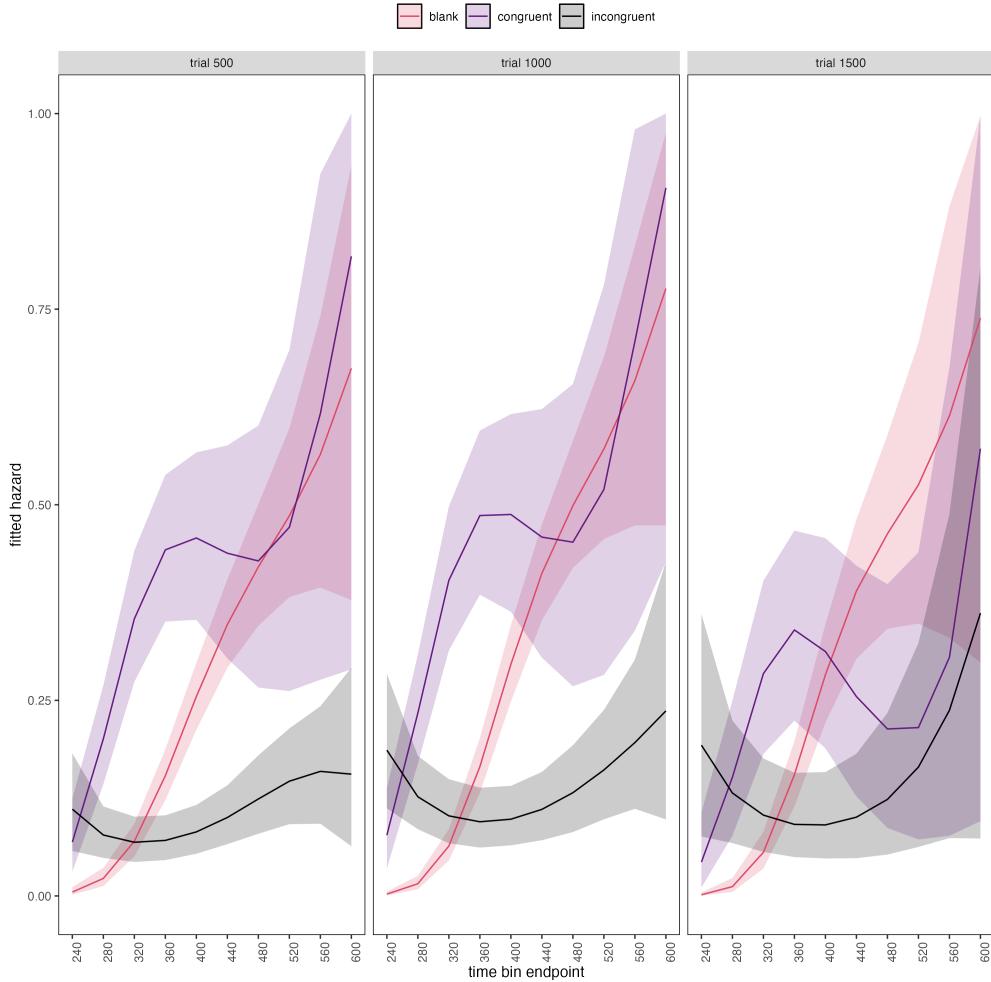
545 Towards the end of the experiment (trial 1500), both congruent and incongruent  
546 primes have positive and negative effects. Positive effects are present in bins (200,240] and  
547 (240,280]. Incongruent primes again have negative effects in bins (320,360], (360,400], and  
548 (400,440], and congruent primes now also have negative effects in bins (360,400] and  
549 (400,440].

550 These results show that the effect of prime-target congruency changes not only on the  
551 across-bin/within-trial time scale (variable period\_9), but also on the  
552 across-trial/within-experiment time scale (variable trial\_c). The fact that congruent  
553 primes generate negative effects for 80 ms (compared to no-prime trials) towards the end of

554 the experiment, while incongruent primes generate negative effects for 120 ms throughout  
555 the experiment, strongly suggests the involvement of separate cognitive processes.

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

572 To conclude this Tutorial 2a, Figure 7 shows the model-based hazard functions for  
573 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.

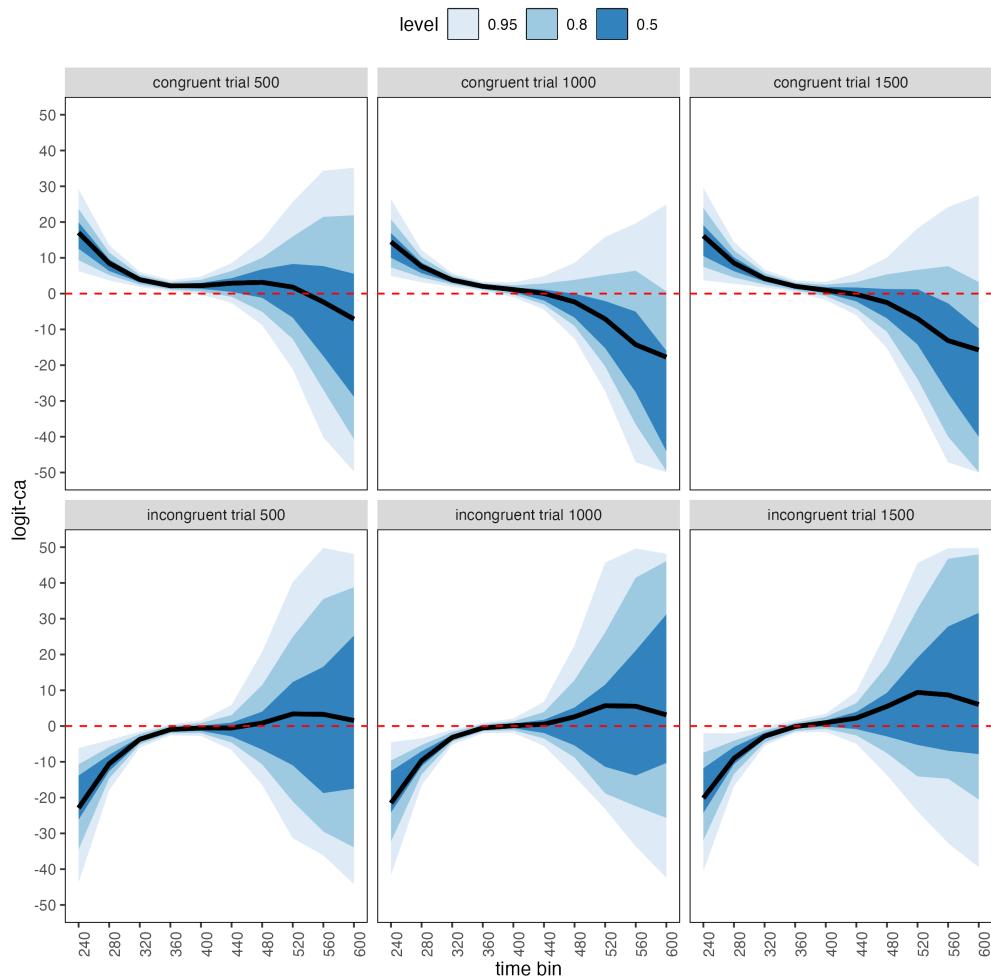
574 **4.4 Tutorial 2b: Fitting Bayesian conditional accuracy models**

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

581 To make inferences from the parameter estimates in model M4\_ca, we summarize the

582 draws from the posterior distributions of the effects of congruent and incongruent primes  
 583 relative to the blank prime condition, in each time bin for trial numbers 500, 1000, and  
 584 1500, in terms of point and interval estimates.

585 Figure 8 shows one point (mean) and three highest posterior density interval  
 586 (50/80/95%) estimates for the effects of congruent and incongruent primes relative to  
 587 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.

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

589 effects of congruent and incongruent primes relative to the blank prime condition in trials  
 590 500, 1000, and 1500, in terms of a point estimate (the mean) and the upper and lower  
 591 bounds of the 95% highest posterior density interval. Also, by exponentiating the mean we  
 592 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

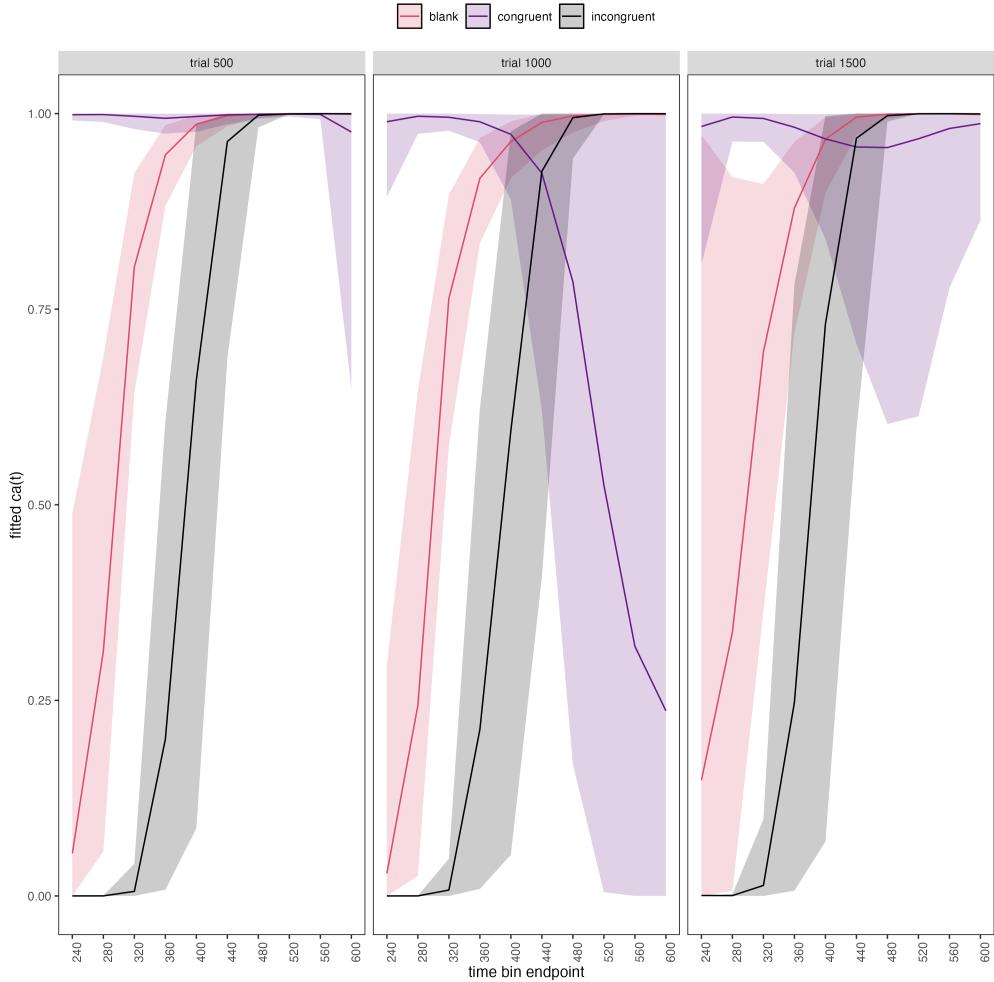
bin_endpoint	condition	mean	.lower	.upper	.width	odds ratio
240	i500	-23.34	-44.42	-4.87	0.95	0.0000000000730
280	i500	-10.55	-17.93	-3.94	0.95	0.0000261643171
320	i500	-3.71	-6.06	-1.51	0.95	0.0246001196011
360	i500	-0.97	-2.57	0.56	0.95	0.3780876098423
400	i500	-0.52	-2.75	1.55	0.95	0.5916606671310
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.

593

594       Based on Figure 8 and Table 5, we see that throughout the experiment (trials 500,  
595   1000, and 1500), congruent primes have a positive effect on logit-ca(t) in time bins  
596   (200,240], (240,280], (280,320], and (320,360], relative to the logit-ca(t) estimates in the  
597   baseline condition (no prime; red striped lines in Figure 8). For example, the odds ratio for  
598   congruent primes in bin (320,360] in trial 500 shows that the odds of a correct response are  
599   estimated to be 8.89 times higher than the odds of a correct response when there is no  
600   prime. Incongruent primes have a negative effect on logit-ca(t) in time bins (200,240],  
601   (240,280], and (280,320] throughout the experiment, relative to the logit-ca(t) estimates in  
602   the baseline condition (no prime; red striped lines).

603       To conclude this Tutorial 2b, Figure 9 shows the model-based ca(t) functions for each  
604   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.

#### 605 4.5 Tutorial 3a: Fitting Frequentist hazard models

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

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

613 of model M3\_f from glmer().

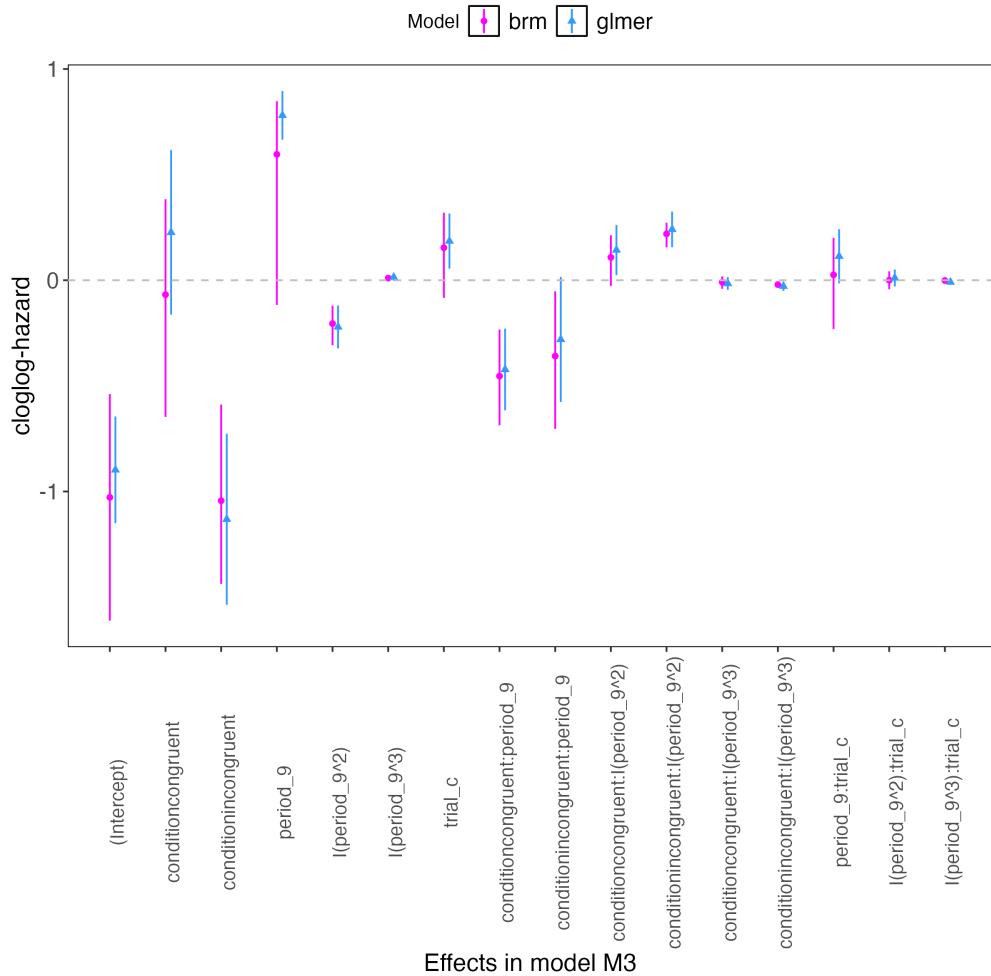


Figure 10. Parameter estimates for model M3 from brm() – means and 95% quantile intervals – and model M3\_f from glmer() – maximum likelihood estimates and 95% confidence intervals.

614 Figure 10 confirms that the parameter estimates from both Bayesian and frequentist  
 615 models are pretty similar. However, the random effects structure of model M3\_f was  
 616 already too complex for the frequentist model as it did not converge and resulted in a  
 617 singular fit. This is of course one of the reasons why Bayesian modeling has become so  
 618 popular in recent years. But the price you pay for being able to fit more complex models in  
 619 a Bayesian framework is computation time. In other words, as we have noted throughout,

620 some of the Bayesian models in Tutorials 2a and 2b took several hours to build.

621 **4.6 Tutorial 3b: Fitting Frequentist conditional accuracy models**

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

628 **5. Discussion**

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

**643 5.1 Advantages of hazard analysis**

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

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

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

669 which are typically treated as outliers and are discarded before calculating a mean. This  
670 orthodox procedure can lead to a sampling bias, however, which results in underestimation  
671 of the mean. By introducing a fixed censoring time for all trials at the end of the analysis  
672 time window, trials with long RTs are not discarded but contribute to the risk set of each  
673 bin.

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

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

687 **5.2 Model complexity versus interpretability**

688 Models for discrete-time  $h(t)$  and  $ca(t)$  can quickly become very complex when adding  
689 more than 1 time scale, due to the many possible higher-order interactions. For example,  
690 model M4 contains two time scales as covariates: the passage of time on the across-bin or  
691 within-trial time scale (variable period\_9), and the passage of time on the across-trial or  
692 within-experiment time scale (variable trial\_c). However, when trials are presented in  
693 blocks, and blocks of trials within sessions, and when the experiment comprises three

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

702 **5.3 Individual differences**

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

716 But even when using a response deadline, participants can differ qualitatively in the  
717 effects they display (see Panis, 2020). One way to deal with this is to describe and  
718 interpret the different patterns. Another way is to run a clustering algorithm on the  
719 individual hazard estimates across all conditions. The obtained dendrogram can then be

720 used to identify a (hopefully big) cluster of participants that behave similarly, and to  
721 identify a (hopefully small) cluster of participants with outlying behavioral patterns. One  
722 might then exclude the outlying participants before fitting a hazard model.

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

729 **5.4 Limitation(s)**

730 Compared to the orthodox method – comparing mean-averages between conditions –  
731 the most important limitation of multilevel hazard and conditional accuracy modeling is  
732 that it might take a long time to estimate the parameters using Bayesian methods or the  
733 model might have to be simplified significantly to use frequentist methods.

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

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

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

754 **5.5 Extensions**

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

770

## 6. Conclusions

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

782

## References

- 783 Allison, P. D. (1982). Discrete-Time Methods for the Analysis of Event Histories.  
784       *Sociological Methodology*, 13, 61. <https://doi.org/10.2307/270718>
- 785 Allison, P. D. (2010). *Survival analysis using SAS: A practical guide* (2. ed). Cary, NC:  
786       SAS Press.
- 787 Aust, F. (2019). *Citr: 'RStudio' add-in to insert markdown citations*. Retrieved from  
788       <https://github.com/crsh/citr>
- 789 Aust, F., & Barth, M. (2023). *papaja: Prepare reproducible APA journal articles with R*  
790       *Markdown*. Retrieved from <https://github.com/crsh/papaja>
- 791 Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., &  
792       Andrews, T. J. (2021). Power contours: Optimising sample size and precision in  
793       experimental psychology and human neuroscience. *Psychological Methods*, 26(3),  
794       295–314. <https://doi.org/10.1037/met0000337>
- 795 Barth, M. (2023). *tinylabes: Lightweight variable labels*. Retrieved from  
796       <https://cran.r-project.org/package=tinylabes>
- 797 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects  
798       models using lme4. *Journal of Statistical Software*, 67(1), 1–48.  
799       <https://doi.org/10.18637/jss.v067.i01>
- 800 Bates, D., Maechler, M., & Jagan, M. (2024). *Matrix: Sparse and dense matrix classes and*  
801       *methods*. Retrieved from <https://CRAN.R-project.org/package=Matrix>
- 802 Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in r  
803       using futures. *The R Journal*, 13(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>
- 804 Blossfeld, H.-P., & Rohwer, G. (2002). *Techniques of event history modeling: New*  
805       *approaches to causal analysis*, 2nd ed (pp. x, 310). Mahwah, NJ, US: Lawrence  
806       Erlbaum Associates Publishers.
- 807 Box-Steffensmeier, J. M. (2004). Event history modeling: A guide for social scientists.  
808       Cambridge: University Press.

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

- 836 population prevalence. *eLife*, 10, e62461. <https://doi.org/10.7554/eLife.62461>
- 837 Kantowitz, B. H., & Pachella, R. G. (2021). The Interpretation of Reaction Time in  
838 Information-Processing Research 1. *Human Information Processing*, 41–82.  
839 <https://doi.org/10.4324/9781003176688-2>
- 840 Kay, M. (2023). *tidybayes: Tidy data and geoms for Bayesian models*.  
841 <https://doi.org/10.5281/zenodo.1308151>
- 842 Kelso, J. A. S., Dumas, G., & Tognoli, E. (2013). Outline of a general theory of behavior  
843 and brain coordination. *Neural Networks: The Official Journal of the International*  
844 *Neural Network Society*, 37, 120–131. <https://doi.org/10.1016/j.neunet.2012.09.003>
- 845 Kurz, A. S. (2023a). *Applied longitudinal data analysis in brms and the tidyverse* (version  
846 0.0.3). Retrieved from <https://bookdown.org/content/4253/>
- 847 Kurz, A. S. (2023b). *Statistical rethinking with brms, ggplot2, and the tidyverse: Second*  
848 *edition* (version 0.4.0). Retrieved from <https://bookdown.org/content/4857/>
- 849 Landes, J., Engelhardt, S. C., & Pelletier, F. (2020). An introduction to event history  
850 analyses for ecologists. *Ecosphere*, 11(10), e03238. <https://doi.org/10.1002/ecs2.3238>
- 851 Luce, R. D. (1991). *Response times: Their role in inferring elementary mental organization*  
852 (1. issued as paperback). Oxford: Univ. Press.
- 853 McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and*  
854 *Stan* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315372495>
- 855 Meyer, D. E., Osman, A. M., Irwin, D. E., & Yantis, S. (1988). Modern mental  
856 chronometry. *Biological Psychology*, 26(1-3), 3–67.  
857 [https://doi.org/10.1016/0301-0511\(88\)90013-0](https://doi.org/10.1016/0301-0511(88)90013-0)
- 858 Müller, K., & Wickham, H. (2023). *Tibble: Simple data frames*. Retrieved from  
859 <https://CRAN.R-project.org/package=tibble>
- 860 Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes*. Retrieved from  
861 <https://CRAN.R-project.org/package=RColorBrewer>
- 862 Panis, S. (2020). How can we learn what attention is? Response gating via multiple direct

- 863 routes kept in check by inhibitory control processes. *Open Psychology*, 2(1), 238–279.
- 864 <https://doi.org/10.1515/psych-2020-0107>
- 865 Panis, S., Moran, R., Wolkersdorfer, M. P., & Schmidt, T. (2020). Studying the dynamics  
866 of visual search behavior using RT hazard and micro-level speed–accuracy tradeoff  
867 functions: A role for recurrent object recognition and cognitive control processes.  
868 *Attention, Perception, & Psychophysics*, 82(2), 689–714.  
869 <https://doi.org/10.3758/s13414-019-01897-z>
- 870 Panis, S., Schmidt, F., Wolkersdorfer, M. P., & Schmidt, T. (2020). Analyzing Response  
871 Times and Other Types of Time-to-Event Data Using Event History Analysis: A Tool  
872 for Mental Chronometry and Cognitive Psychophysiology. *I-Perception*, 11(6),  
873 2041669520978673. <https://doi.org/10.1177/2041669520978673>
- 874 Panis, S., & Schmidt, T. (2016). What Is Shaping RT and Accuracy Distributions? Active  
875 and Selective Response Inhibition Causes the Negative Compatibility Effect. *Journal of*  
876 *Cognitive Neuroscience*, 28(11), 1651–1671. [https://doi.org/10.1162/jocn\\_a\\_00998](https://doi.org/10.1162/jocn_a_00998)
- 877 Panis, S., & Schmidt, T. (2022). When does “inhibition of return” occur in spatial cueing  
878 tasks? Temporally disentangling multiple cue-triggered effects using response history  
879 and conditional accuracy analyses. *Open Psychology*, 4(1), 84–114.  
880 <https://doi.org/10.1515/psych-2022-0005>
- 881 Panis, S., Torfs, K., Gillebert, C. R., Wagemans, J., & Humphreys, G. W. (2017).  
882 Neuropsychological evidence for the temporal dynamics of category-specific naming.  
883 *Visual Cognition*, 25(1-3), 79–99. <https://doi.org/10.1080/13506285.2017.1330790>
- 884 Panis, S., & Wagemans, J. (2009). Time-course contingencies in perceptual organization  
885 and identification of fragmented object outlines. *Journal of Experimental Psychology:*  
886 *Human Perception and Performance*, 35(3), 661–687.  
887 <https://doi.org/10.1037/a0013547>
- 888 Pedersen, T. L. (2024). *Patchwork: The composer of plots*. Retrieved from  
889 <https://CRAN.R-project.org/package=patchwork>

- 890 Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in s and s-PLUS*. New York:  
891 Springer. <https://doi.org/10.1007/b98882>
- 892 R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna,  
893 Austria: R Foundation for Statistical Computing. Retrieved from  
894 <https://www.R-project.org/>
- 895 Ripley, B., Venables, B., Bates, D. M., ca 1998), K. H. (partial. port, ca 1998), A. G.  
896 (partial. port, & polr), D. F. (support. functions for. (2024). *MASS: Support Functions*  
897 and *Datasets for Venables and Ripley's MASS*.
- 898 Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling  
899 Change and Event Occurrence*. Oxford, New York: Oxford University Press.
- 900 Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design.  
901 *Psychonomic Bulletin & Review*, 25(6), 2083–2101.  
902 <https://doi.org/10.3758/s13423-018-1451-8>
- 903 Steele, F., Goldstein, H., & Browne, W. (2004). A general multilevel multistate competing  
904 risks model for event history data, with an application to a study of contraceptive use  
905 dynamics. *Statistical Modelling*, 4(2), 145–159.  
906 <https://doi.org/10.1191/1471082X04st069oa>
- 907 Teachman, J. D. (1983). Analyzing social processes: Life tables and proportional hazards  
908 models. *Social Science Research*, 12(3), 263–301.  
909 [https://doi.org/10.1016/0049-089X\(83\)90015-7](https://doi.org/10.1016/0049-089X(83)90015-7)
- 910 Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical  
911 distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, 108(3),  
912 551–567. <https://doi.org/10.1037/0033-2909.108.3.551>
- 913 Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics.  
914 *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)
- 915 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New  
916 York. Retrieved from <https://ggplot2.tidyverse.org>

- 917 Wickham, H. (2023a). *Forcats: Tools for working with categorical variables (factors)*.  
918 Retrieved from <https://forcats.tidyverse.org/>
- 919 Wickham, H. (2023b). *Stringr: Simple, consistent wrappers for common string operations*.  
920 Retrieved from <https://stringr.tidyverse.org>
- 921 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ...  
922 Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43),  
923 1686. <https://doi.org/10.21105/joss.01686>
- 924 Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A  
925 grammar of data manipulation*. Retrieved from <https://dplyr.tidyverse.org>
- 926 Wickham, H., & Henry, L. (2023). *Purrr: Functional programming tools*. Retrieved from  
927 [https://purrr.tidyverse.org/](https://purrr.tidyverse.org)
- 928 Wickham, H., Hester, J., & Bryan, J. (2024). *Readr: Read rectangular text data*. Retrieved  
929 from <https://readr.tidyverse.org>
- 930 Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data*. Retrieved from  
931 <https://tidyr.tidyverse.org>
- 932 William Matthew Makeham. (1860). *On the Law of Mortality and the Construction of  
933 Annuity Tables*. The Assurance Magazine, and Journal of the Institute of Actuaries.

934

**Supplementary material**

935 **A. Definitions of discrete-time hazard, survivor, and conditional accuracy  
936 functions**

937 The shape of a distribution of waiting times can be described in multiple ways (Luce,  
938 1991). After dividing time in discrete, contiguous time bins indexed by  $t$ , let  $RT$  be a  
939 discrete random variable denoting the rank of the time bin in which a particular person's  
940 response occurs in a particular experimental condition. Discrete-time EHA focuses on the  
941 discrete-time hazard function

942 
$$h(t) = P(RT = t | RT \geq t) \quad (1)$$

943 and the discrete-time survivor function

944 
$$S(t) = P(RT > t) = [1-h(t)].[1-h(t-1)].[1-h(t-2)] \dots [1-h(1)] \quad (2)$$

945 and not on the probability mass function

946 
$$P(t) = P(RT = t) = h(t).S(t-1) \quad (3)$$

947 nor the cumulative distribution function

948 
$$F(t) = P(RT \leq t) = 1-S(t) \quad (4)$$

949 The discrete-time hazard function of event occurrence gives you the probability that  
950 the event occurs (sometime) in bin  $t$ , given that the event has not occurred yet in previous  
951 bins. While the discrete-time hazard function assesses the unique risk of event occurrence  
952 associated with each time bin, the discrete-time survivor function cumulates the bin-by-bin  
953 risks of event *nonoccurrence* to obtain the probability that the event occurs after bin  $t$ . The  
954 probability mass function cumulates the risk of event occurrence in bin  $t$  with the risks of  
955 event nonoccurrence in bins 1 to  $t-1$ . From equation 3 we find that hazard in bin  $t$  is equal  
956 to  $P(t)/S(t-1)$ .

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

958 discrete-time conditional accuracy function

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

960 which gives you the probability that a response is correct given that it is emitted in time

961 bin  $t$  (Allison, 2010; Kantowitz & Pachella, 2021; Wickelgren, 1977). This latter function is

962 also known as the micro-level speed-accuracy tradeoff (SAT) function.

963 The survivor function provides a context for the hazard function, as  $S(t-1) = P(RT >$

964  $t-1) = P(RT \geq t)$  tells you on how many percent of the trials the estimate  $h(t) = P(RT =$

965  $t \mid RT \geq t$ ) is based. The probability mass function provides a context for the conditional

966 accuracy function, as  $P(t) = P(RT = t)$  tells you on how many percent of the trials the

967 estimate  $ca(t) = P(\text{correct} \mid RT = t)$  is based.

968 When time is treated as a continuous variable, let  $RT$  be a continuous random variable

969 denoting a particular person's response time in a particular experimental condition.

970 Because waiting times can only increase, continuous-time EHA does not focus on the

971 cumulative distribution function  $F(t) = P(RT \leq t)$  and its derivative, the probability

972 density function  $f(t) = F(t)'$ , but on the survivor function  $S(t) = P(RT > t)$  and the

973 hazard rate function  $\lambda(t) = f(t)/S(t)$ . The hazard rate function gives you the instantaneous

974 *rate* of event occurrence at time point  $t$ , given that the event has not occurred yet.

## 975 B. Custom functions for descriptive discrete-time hazard analysis

976 We defined 13 custom functions that we list here.

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

- 981     • setup\_lt\_2IV(ptb) : set up a life table for each combination of levels of 2  
982                 independent variables
- 983     • calc\_ca(df) : estimate the conditinal accuracies when there is 1 independent variable
- 984     • calc\_ca\_2IV(df) : estimate the conditional accuraiies when there are 2 independent  
985                 variables
- 986     • join\_lt\_ca(df1,df2) : add the ca(t) estimates to the life tables (1 independent  
987                 variable)
- 988     • join\_lt\_ca\_2IV(df1,df2) : add the ca(t) estimates to the life tables (2 independent  
989                 variables)
- 990     • extract\_median(df) : estimate quantiles S(t).50 (1 independent variable)
- 991     • extract\_median\_2IV(df) : estimate quantiles S(t).50 (2 independent variables)
- 992     • plot\_eha(df,subj,haz\_yaxis) : create plots of the discrete-time functions (1  
993                 independent variable)
- 994     • plot\_eha\_2IV(df,subj,haz\_yaxis) : create plots of the discrete-time functions (2  
995                 independent variables)
- 996     • plot\_eha\_agg(df,subj,haz\_yaxis) : create 1 plot for aggregated data (1 independent  
997                 variable)

998     When you want to analyse simple RT data from a detection experiment with one  
999     independent variable, the functions calc\_ca() and join\_lt\_ca() should not be used, and  
1000    the code to plot the conditional accuracy functions should be removed from the function  
1001    plot\_eha(). When you want to analyse simple RT data from a detection experiment with  
1002    two independent variables, the functions calc\_ca\_2IV() and join\_lt\_ca\_2IV() should not  
1003    be used, and the code to plot the conditional accuracy functions should be removed from  
1004    the function plot\_eha\_2IV().

1005 **C. Link functions**

1006 Popular link functions include the logit link and the complementary log-log link, as  
 1007 shown in Figure 11.

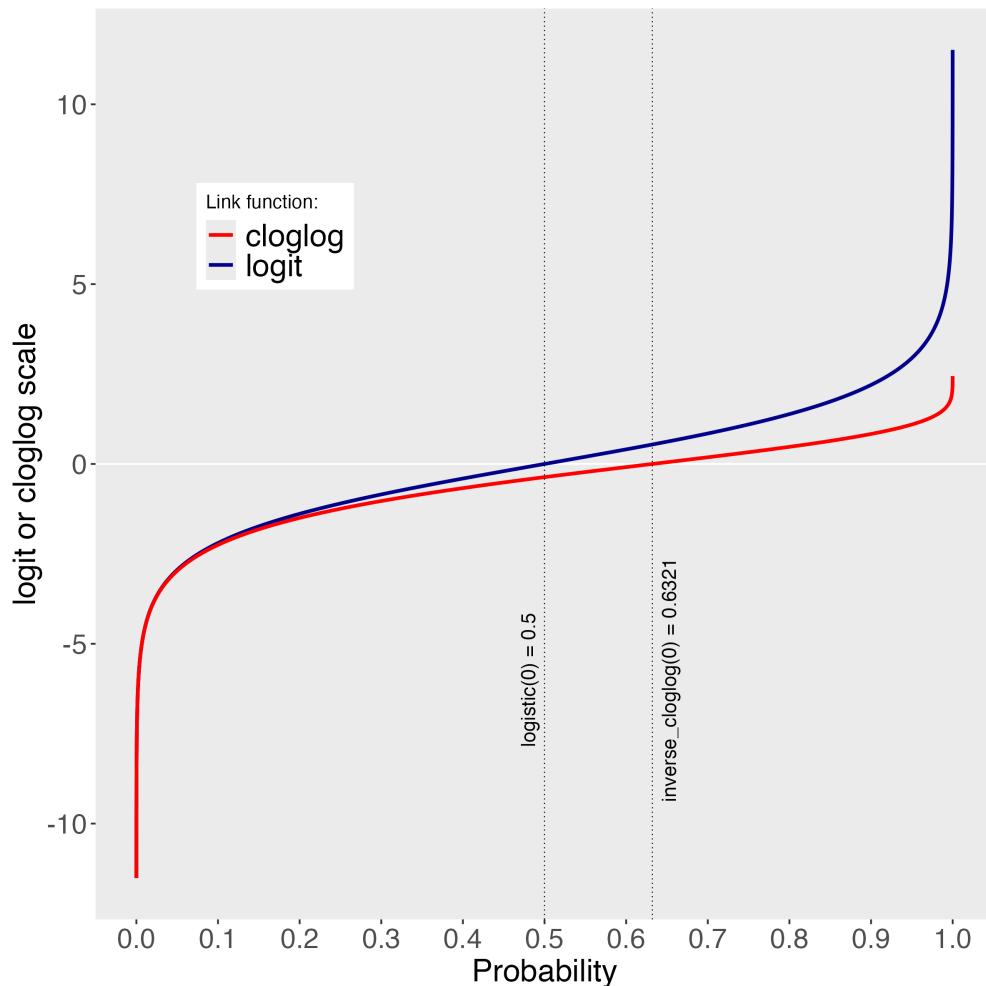


Figure 11. The logit and cloglog link functions.

1008 **D. Regression equations**

1009 An example (single-level) discrete-time hazard model with three predictors (TIME,  
 1010 X<sub>1</sub>, X<sub>2</sub>), the cloglog link function, and a third-order polynomial specification for TIME can  
 1011 be written as follows:

$$\begin{aligned} 1012 \quad & \text{cloglog}[h(t)] = \ln(-\ln[1-h(t)]) = [\alpha_1 \text{ONE} + \alpha_2(\text{TIME}-9) + \alpha_3(\text{TIME}-9)^2] + [\beta_1 X_1 + \\ 1013 \quad & \beta_2 X_2 + \beta_3 X_2(\text{TIME}-9)] \end{aligned}$$

1014        The main predictor variable TIME is the time bin index  $t$  that is centered on value 9  
 1015      in this example. The first set of terms within brackets, the alpha parameters multiplied by  
 1016      their polynomial specifications of (centered) time, represents the shape of the baseline  
 1017      cloglog-hazard function (i.e., when all predictors  $X_i$  take on a value of zero). The second set  
 1018      of terms (the beta parameters) represents the vertical shift in the baseline cloglog-hazard  
 1019      for a 1 unit increase in the respective predictor variable. Predictors can be discrete,  
 1020      continuous, and time-varying or time-invariant. For example, the effect of a 1 unit increase  
 1021      in  $X_1$  is to vertically shift the whole baseline cloglog-hazard function by  $\beta_1$  cloglog-hazard  
 1022      units. However, if the predictor interacts linearly with TIME (see  $X_2$  in the example), then  
 1023      the effect of a 1 unit increase in  $X_2$  is to vertically shift the predicted cloglog-hazard in bin  
 1024      9 by  $\beta_2$  cloglog-hazard units (when  $\text{TIME}-9 = 0$ ), in bin 10 by  $\beta_2 + \beta_3$  cloglog-hazard  
 1025      units (when  $\text{TIME}-9 = 1$ ), and so forth. To interpret the effects of a predictor, its  $\beta$   
 1026      parameter is exponentiated, resulting in a hazard ratio (due to the use of the cloglog link).  
 1027      When using the logit link, exponentiating a  $\beta$  parameter results in an odds ratio.

1028        An example (single-level) discrete-time hazard model with a general specification for  
 1029      TIME (separate intercepts for each of six bins, where D1 to D6 are binary variables  
 1030      identifying each bin) and a single predictor ( $X_1$ ) can be written as follows:

$$\begin{aligned} 1031 \quad & \text{cloglog}[h(t)] = \ln(-\ln[1-h(t)]) = [\alpha_1 D1 + \alpha_2 D2 + \alpha_3 D3 + \alpha_4 D4 + \alpha_5 D5 + \alpha_6 D6] + \\ 1032 \quad & [\beta_1 X_1] \end{aligned}$$

1033 **E. Prior distributions**

1034        To gain a sense of what prior *logit* values would approximate a uniform distribution  
 1035      on the probability scale, Kurz (2023a) simulated a large number of draws from the  
 1036      Uniform(0,1) distribution, converted those draws to the log-odds metric, and fitted a

1037 Student's t distribution. Row C in Figure 12 shows that using a t-distribution with 7.61  
 1038 degrees of freedom and a scale parameter of 1.57 as a prior on the logit scale, approximates  
 1039 a uniform distribution on the probability scale. According to Kurz (2023a), such a prior  
 1040 might be a good prior for the intercept(s) in a logit-hazard model, while the  $N(0,1)$  prior in  
 1041 row D might be a good prior for the non-intercept parameters in a logit-hazard model, as it  
 1042 gently regularizes  $p$  towards .5 (i.e., a zero effect on the logit scale).

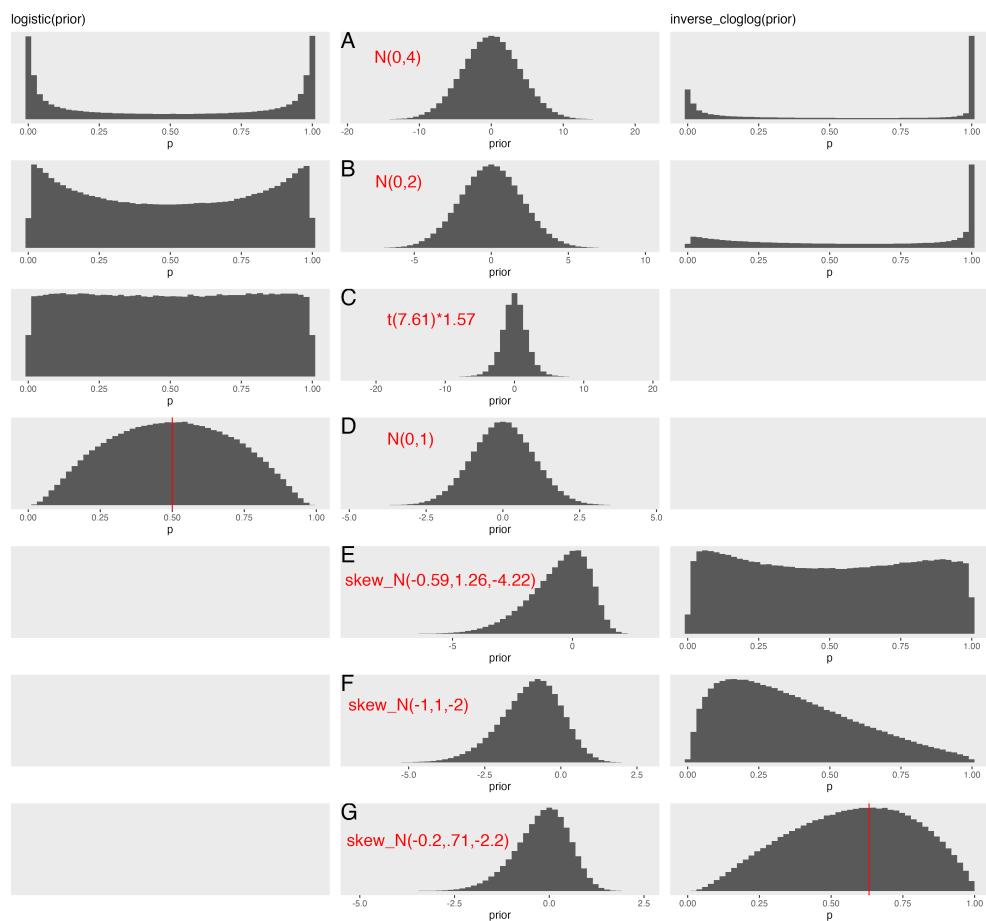


Figure 12. Prior distributions on the logit and/or cloglog scales (middle column), and their implications on the probability scale after applying the inverse-logit (or logistic) transformation (left column), and the inverse-cloglog transformation (right column).

1043 To gain a sense of what prior *cloglog* values would approximate a uniform distribution  
 1044 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).