What, me saccade?

Abstract

Basically there is the VWP and it is used as a proxy measure for word recognition. This use follows from Allopenna 1996, in which he showed the the proportion of fixations to different referents matches what would be expected with TRACE (after suitable transformation, of course). This resulted in over two decades of VWP use for such purposes. In 2021, McMurray asked if that curve really was what we thought it was. Through an analysis of generating hypotheses of increasing (but still relatively minimal) complexity via simulation, bob showed that even in cases of moderate complexity, we were not able to positively recover the generating curve responsible for eye movements. why was this, and what are the implications for understanding these curves? In this paper, we revisit the 2021 princess bride paper and offer an explanation for the demonstrated bias of the simulations. from this, we propose a new method for using VWP data to estimate underlying activation curves. We conclude by demonstrating consistency of the proposed method with existing continuous mapping models of word activation.

Notes:

1. Sections that are more “narrative” are less fleshed out. This includes VWP, TRACE, general history, etc., Some of these sections I just said what I would say
2. Citations are hard coded in here awaiting a bib to be created
3. Some plots/graphics need to be redone for size
4. There is some meta commentary, partially for the reader, mostly for me

1 Introduction

Spoken words create analog signals that are processed by the brain in real time. That is, as the spoken word unfolds, a cohort of possible resolutions are considered until the target word is recognized. The degree to which a particular candidate word is recognized is known as activation. An important part of this process involves not only correctly identifying the word but also eliminating competitors. For example, we might consider a discrete unfolding of the word “elephant” as “el-e-phant”. At the onset of “el”, a listener may activate a cohort of potential resolutions such as “elephant”, “electricity”, or “elder”, all of which may be considered competitors. With the subsequent “el-e”, words consistent with the received signal, such as “elephant” and “electricity” remain active competitors, while incompatible words, such as “elder”, are eliminated. Such is a rough description of this process, continuing until the ambiguity is resolved and a single word remains.

Our interest is in measuring the degree of activation of a target, relative to competitors. Activation, however, is not measured directly, and we instead rely on what can be observed with eye-tracking data, collected in the context of the Visual World Paradigm (VWP) (Tannenhaus 1995). In the last few years, researchers have begun to reexamine some of the underlying assumptions associated with the VWP, calling into question the validity or interpretation of current methods. We present here a brief history of word recognition in the context of the VWP, along with an examination of contemporary concerns. We address some of these concerns directly, presenting an alternate method for relating eye-tracking data to lexical activation.

This section needs work but it mostly covers the gist of what I am trying to convey, namely we are about to go from history → current state of the world → proposal and comparison → results.

2 A brief history

We begin with a brief history to give context to later discussion. In particular, we will consider one of the leading theoretical models in speech perception, TRACE, followed by the introduction of the leading experimental paradigm, the VWP. We examine empirical evidence for the relation between these, and relevant theoretical advancements that have been made. Topics here are presented only briefly and limited to those directly relevant to the present work. For a fuller discussion of the history and uses of VWP, use google. (Or Huettig 2011b?)

An outline of the presentation (for internal use only):

1. VWP by Tannenhaus 1995
2. VWP + TRACE, Allopenna 1996 (trace aspect no longer relevant, just an aside)
3. As far as I can tell, it’s Bob’s 2010 paper that was among first to
   (a) Look at individual differences in word recognition (not counting the ortho polynomial fits) (also relevant for the “group distribution of curves” hypothesis) and
   (b) Introduce parametric forms to be fit to the data (the assumption we continue to run with), or at very least, introduce ones that are interpretable

All of the paragraphs in this section are narrative and not mission critical. Need to be fleshed out

VWP To briefly illustrate, the VWP is an experimental design in which participants undergo a series of trials to identify a spoken word. Typically, each trial has a single target word, along with multiple competitors. The target word is spoken, and participants are asked to identify and select an image on screen associated with the spoken word. Eye movements and fixations are recorded as this process unfolds, with the location of the participants' eyes serving as proxy for which words/images are being considered.

Proportion of fixation born It was against simulated TRACE data that Allopenna (1998) found a tractable way of analyzing eye tracking data. By coding the period of a fixation as a 0 or 1 for each referent and taking the average of fixations towards a referent at each time point, Allopenna was able to create a "fixation proportion" curve that largely reflected the shape and competitive dynamics of word activation suggested by TRACE, both for the target object, as well as competitors. This also served to establish a simple linking hypothesis, specifically, "We made the general assumption that the probability of initiating an eye movement to fixate on a target object o at time t is a direct function of the probability that o is the target given the speech input and where the probability of fixating o is determined by the activation level of its lexical entry relative to the activation of other potential targets." Further of note is what this linking hypothesis does not include, namely:

1. No assumption that scanning patterns in and of themselves reveal underlying cognitive processes
2. No assumption that the fixation location at time t necessarily reveals where attention is directed (only probabilistically related to attention)

Other assumptions included here include that language processing proceeds independent of vision (Magnuson 2019), and that visual objects are not automatically activated. Or, more succinctly, it assumes that fixation proportions over time provide an essentially direct index of lexical activation, whereby the probability of fixating an object increases as the likelihood that it has been referred to increases.

While other linking hypotheses have been presented (Magnuson 2019), that there is *some* link between the function of fixation proportions and activation has guided the last 25 years of VWP research.

Parametric Methods and Individual Curves While there have most certainly been advancements to the use of the VWP for speech perception and recognition (and expanded into related domains, such as sentence processing and characterizing language disorders (according to Bob)), we limit ourselves here to one in particular. In 2010, McMurray et al expanded the domain of the VWP by introducing emphasis on individual differences in participant activation curves. Two aspects of this paper are relevant here. First, although they were not the first to introduce non-linear functions to be fit to observed data, they did introduce a number of important parametric functions in use today, namely the four (or five) parameter logistic and the double-gauss (asymmetrical gauss), the primary benefit being that the parameters of these functions are interpretable, that is, they "describe readily observable properties." Second, which I suppose was also introduced by Mirman (2008) to some degree (though I have not read it yet, just pulling from Bob) is specifying individual subject curves across participants. This has been critical in that:

1. The parameters of the functions describe interpretable properties
2. This made the idea of distributions of parameters for a particular group a relevant construct

Though not stated directly (given it predates bdots by 8 years), this also served as the impetus for investigating group differences in word activation through the use of bootstrapped differences in time series (Oleson 2017) and the subsequent development of the `bdots` software in R for analyzing such differences. (A history of exploring differences in group curves can be found in (Seedorff 2018)).

This brings us to the current day, where the state of things is such that TRACE-validated VWP data is widely used to measure word recognition by collecting data on individual subjects and fitting to them non-linear parametric curves with interpretable parameters. Context in hand, we are now able to introduce some of the main characters of our story, specifically how data in the VWP is understood and used.

3 Where we are now

This section includes the finer points of the VWP, eye tracking data, and how allopena's introduction ties in with bob's parametric proposition.

3.1 anatomy of eye movements

There are three components of eye movements with which we are concerned. The first two, saccades and fixations, are associated with physical mechanics of eye movements; the third, oculomotor delay, is a phenomenon related to the association between cognitive activation and physiological response. We will briefly introduce each of these topics.

Saccades and fixations: Rather than acting in a continuous sweeping motion as our perceived vision might suggest, our eyes themselves move about in a series of short, ballistic movements, followed by brief periods of stagnation. These, respectively, are the saccades and fixations.

The short ballistic movements are known as saccades, periods of between 20ms-60ms (source? more accurate times?) in which they eye is in motion and during which time we are effectively blind. Once in motion, saccades have no ability to change their intended destination. Following the movement itself is a period of stillness known as a fixation, itself made up of a necessary refraction period from the saccade (time?)

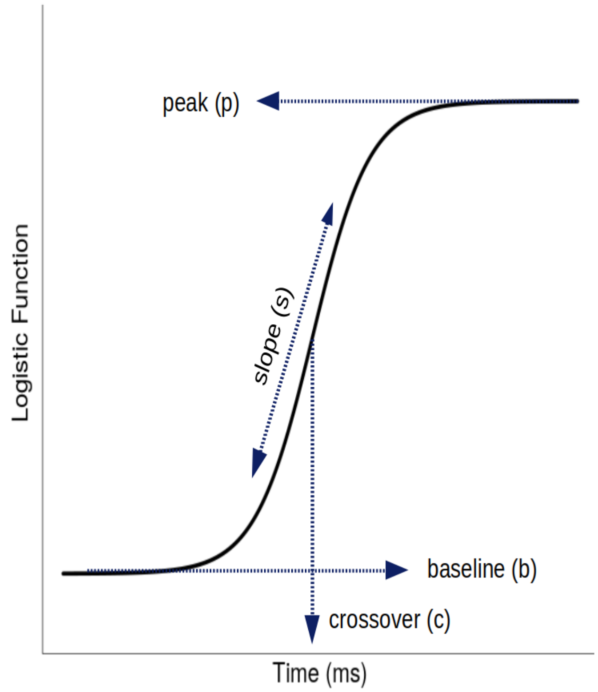


Figure 1: An illustration of the four-parameter logistic and its associated parameters, introduced as a parametric function for fixations to target objects in McMurray 2010. Can describe the parameters in detail, but should also have the formula itself somewhere to be referenced. (Equation **??**)

followed by a period of voluntary fixation; the typical duration of a fixation is (some length). Together, an initiating saccade and its subsequent fixation is known colloquially as a “look”. See Figure **??**.

Oculomotor delay: While the physiological responses are what we can measure, they are not themselves what we are interested in. Rather, we are interested in determining word activation, itself governing the cognitive mechanism facilitating the movements in the eyes. It’s suspected/stated/known (source?) that upon finishing a particular saccade, the mind is already anticipating where it will move next. Length of about 200ms also thrown around a lot. What is relevant for our purpose here, however, is that the period of oculomotor delay is a (likely) random process, resulting in biased observations between what we are able to measure and what we are interested in discovering. How this phenomenon relates to saccades and fixations is demonstrated in Figure **??**.

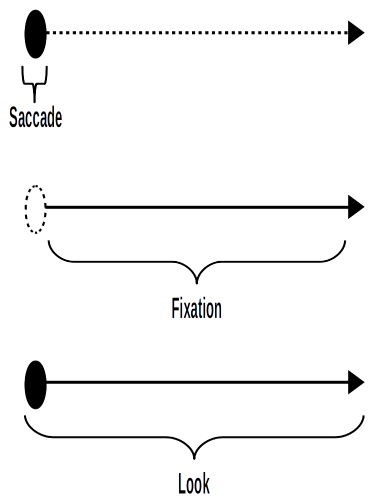


Figure 2: This image needs to be recreated for size. Illustrates saccade, fixation, and look

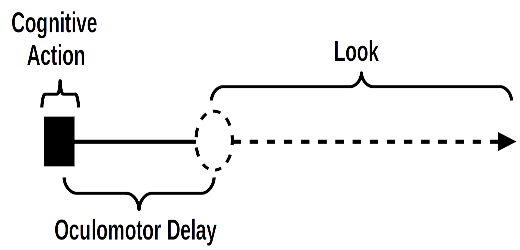


Figure 3: this also could probably be reformatted or made bigger

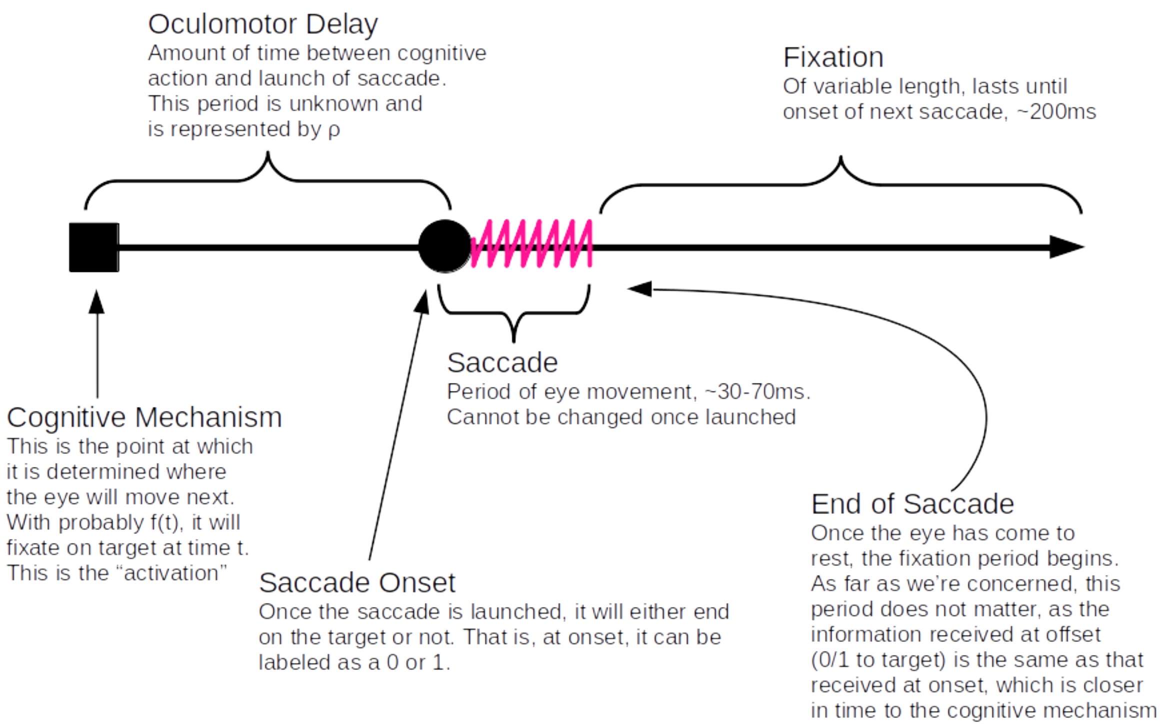


Figure 4: This figure actually doesn't look too bad, but may be better when articulating how saccades measured and why (also includes info on $f(t)$, ρ, etc., so maybe we will present this later around the time of simulation. Mostly here now just to be present

3.2 Activation

What is it, exactly? This would be a good section to introduce notation, specifically that throughout this we will let $f(t)$ be activation in *time*, and in particular $f_\theta(t)$ where

$$f_\theta(t) = \frac{p-b}{1+\exp\left(\frac{4s}{\mathrm{p}-b}(x-t)\right)} + b. \tag{1}$$

Then I can reference Figure **??**. Great, references established.

3.3 VWP data

We now consider how the aforementioned mechanics relate to the VWP. In a typical instantiation of the VWP, a participant is asked to complete a series of trials, during each of which they are presented with a number of competing images on screen (typically four). A verbal cue is given, and the participants are asked to select the image corresponding to the spoken word.

An individual trial of the VWP may be short, lasting anywhere from 1000ms to 2500ms before the correct image is selected. Prior to this, the participants eyes scan the environment, considering images as potential candidates to the spoken word. As this process unfolds, a snapshot of the eye is taken at a series of discrete steps (typically every 4ms) indicating where on the screen the participant is fixated. While there is evidence of cognition happening behind the scenes in a continuous fashion (Spivey, mouse trials), an individual trial of the VWP may contain no more than four to eight total "looks" before the correct image is clicked, resulting in a paucity of data in any given trial.

To create a visual summary of this process aggregated over all of the trials, a la Allopena, a "proportion of fixations" curve is created, aggregating at each discrete time point the average of indicators indicating that a participant is fixated on a particular image. A resulting curve is created for each of the competing categories (target, cohort, rhyme, unrelated), creating an empirical estimate of the activation curve, $f_\theta(t)$. See Figure **??**. Mathematically, it looks like this:

$$y_t = \frac{1}{J}\sum z_{jt} \tag{2}$$

where z_{ijt} is an indicator $\{0, 1\}$ in trial j at time t and such that we have an empirical estimate of the activation curve,

$$f_\theta(t) \equiv y_t. \tag{3}$$

In other words, we see here that it is implicitly assumed that the trajectory of the eye follows the trajectory

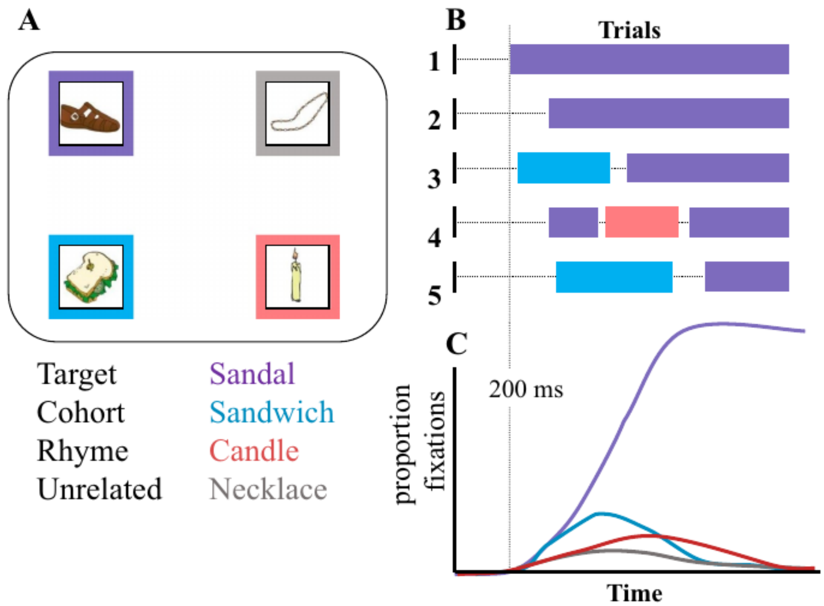


Figure 5: This screenshotted from Bob's princess bridge paper. i would like to reconstruct a similar illustration here as it does a great job illustrating the point. *However*, this section as it stands may make more sense elaborated elsewhere, in particular where I give a mathematical treatment to what the "fixation curve" is

of activation, where the average proportion of fixations at a particular time is a direct estimate of activation. As each individual trial is only made up of a few ballistic movements, the aggregation across trials allows for these otherwise discrete measurements to more closely represent a continuous curve. Curve fitting methods, such as those employed by `bdots`, are then used to construct estimates of function parameters fitted to this curve.

4 Where are we going?

Somewhere I need to be clear with my language in that typically a saccade is a period of movement lasting about 30ms, followed by a fixation lasting however much time. I will be talking specifically about saccade onset or fixation onset which is not associated with a duration but rather a specific instance in time.

Having given due consideration to the state of things are they are, we find ourselves in a time of moral reflection, reexamining the underlying relationship between lexical activation, the mechanism of interest, and the physiological behavior we are able to observe (here, specifically eye-tracking, rather than discussion on other behavioral tasks, i.e., Spivey mouse tracking). This is referred to in the literature as the linking hypothesis. And while there are a number of competing hypothesis, they each share a collection of implicit assumptions relating what is observed to what is being studied.

In particular, we consider a contribution presented by McMurray 2022 in which he probed the relationship between the observed dynamics of the fixations curves and the underlying dynamics of activation under a variety of assumptions. In short, he showed that curves reconstructed using the standard (standard being determing proportion of fixations, may specify that in more detail earlier) analysis in the VWP were poor estimates of the underlying system, with the magnitude of bias increasing on the complexity of the mechanisms involved.

[transition paragraph?]

From allopenna – "We made the general assumption that the probability of initiating an eye movement to fixate on a target object o at time t is a direct function of the probability that o is the target given the speech input and where the probability of fixating o is determined by the activation level of its lexical entry relative to the activation of the other potential targets."

[other transition paragraph?] Really, before I lay out the biases I do have to talk about simulations or at least a generating curve, otherwise it kind of doesn't make any sense. In that case, I should probably just specify the simulation I will do to replicate it in high level detail. Maybe I will introduce the saccade method here in contrast to the standard analysis and THEN describe the sources of bias from the first. That way I can use saccade notation later

From this, and what we ultimately argue here, the observed bias can be partitioned into two distinct components:

1. The first source of bias, which is the primary emphasis of my proposal, is what I call the "added observation" bias. This involves the fact that in a standard analysis of VWP data is, the entire duration of a fixation is indicated with a $\{0, 1\}$ at any time, t, without having observed any behavior associated with the initiation of an eye movement at that time.
2. The second source of bias is "delayed observation bias". This bias arises from the fact that an eye movement launched at some time t was planned at some time prior. This includes both the refractory period of an existing fixation, as well as oculomotor delay

The first source of bias, the "added observation" bias, arises singularly from the fact that a standard analysis does not differentiate between the instance of saccade onset and the subsequent fixations in the observed data. To illustrate, consider a situation in which there is no delayed observation and that a probability that an eye movement launched at time t will fixate on the target is directly determined by activation at time t, a la Allopenna 1998. That is, when we observe a saccade s_t launched at time t, we are sampling directly from the activation curve following some distribution at that point in time,

$$s_t \sim Bin(f_\theta(t)) \tag{4}$$

where $f_\theta(t)$ is assumed to be the activation curve. What to make, then, of the subsequent fixation recorded at $t+1$? Under the current method, the ongoing fixation is treated as a readout of the activation curve at each subsequent time for the duration of the fixation. In other words, we treat the initiation of an eye movement at time t, governed by the underlying dynamics we wish to retrieve, as identical with the subsequent fixation over n time points from $t+1$ to $t+n$, including the period of time in which there is a necessary refractory period and no new information about the underlying activation could possibly be collected from the eye mechanics. An illustration of this bias is given in Figure **??**

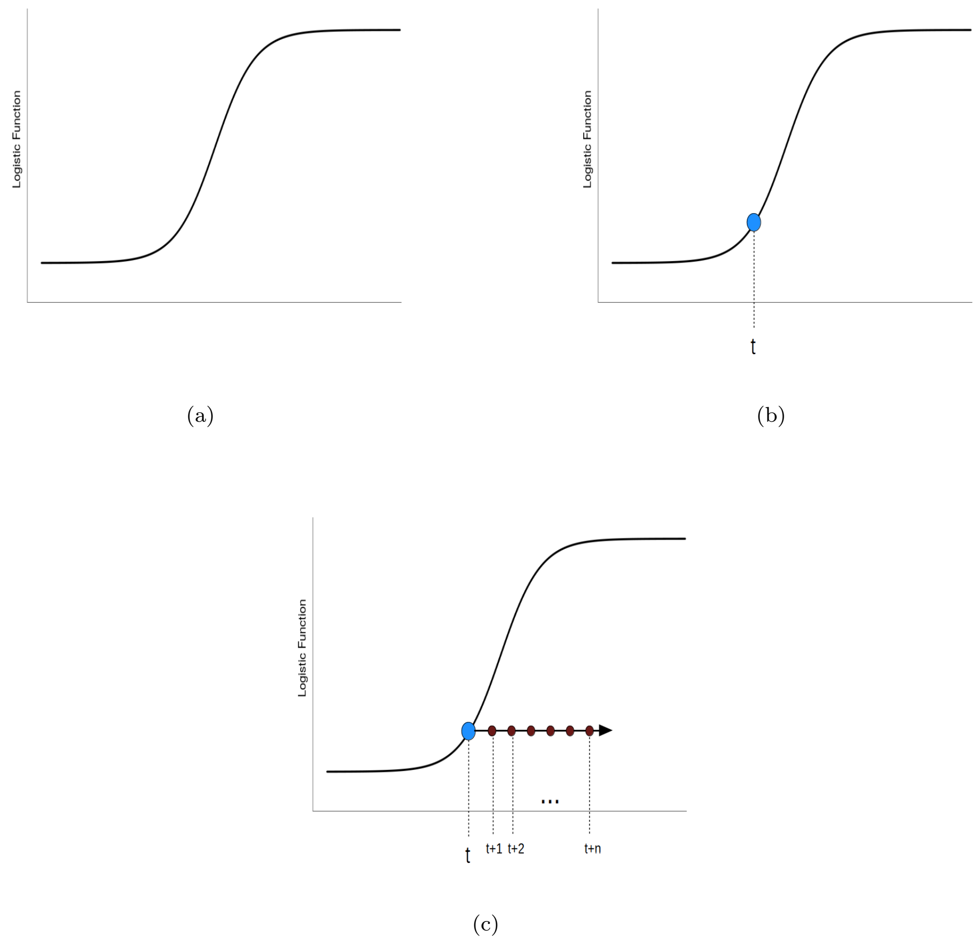


Figure 6: These illustrations can all be made larger (they were made for slides in an image editing program), but they illustrate the main point. **(a.)** here we see an example of a generating logistic function **(b.)** at some time, t, a saccade is launched (in the algorithm, a binomial is drawn with probability $Bin(f_\theta(t))$ **(c.)** at subsequent times, $t+1,\ldots,t+n$, we are recording "observed" data, adding to the proportion of fixations at each time but without having gathered any additional observed data at $f_\theta(t+1),\ldots,f_\theta(t+n)$, thus inflating (or in the case of a monotonically increasing function like the logistic, deflating) the true probability.

The consequence of this is that the *amount* of observed data is artificially inflated. And in the particular case of the four parameter logistic funciton, this acts to artificially *deflate* the observed probability associated with the added observations. That is, as this function is monotonein time, it follows that $f_\theta(t) < f_\theta(t+n)$ for all t and n. As such, a saccade observed at t with some probability $f_\theta(t)$ will also function as an observation at time when the underlying activation is actually $f_\theta(t+n)$, thereby "slowing" the rate of activation. As we will see in the simulations, the result is a delayed crossover parameter and a flatter slope. [comment of relationship of total variation with observed bias, tie back to double gauss]

Finally, a quick comment on the delayed observation bias. It is well established in the literature that it takes around 200ms to plan and launch a saccade meaning that a saccade launched at time t was likely planned around 200ms earlier (Viviani 1990). This phenomenon, known as oculomotor delay, is typically accounted for by shifting the observed data 200ms back before performing any analysis. While this presents no issue when the oculmotor delay is always fixed at 200ms, it is worthwhile considering the impact of this delayed observation when the true delay has an associated variability.

—

While there is no immediate solution to the effects of randomness in the delayed observation bias, we argue that the added observation bias can be rectified by using *only* the observed times associated with saccade onset in the recovery of the underlying dynamics.

Saccade Method: Here are a few points to be made in whatever amount of detail. First, we have to rectify the fact that we are now comparing essentially two different curves: one for the proportion of fixations, the other the probability of launching a saccade. Functionally this may be of little importance. Next, we should mention that we can fit this to the same curve (four parameter logistic) using the exact same methods (bdots). Lastly, we can maybe repeat (or move here) a mathematical description of the saccade method, namely what was shown in Equation **??**. This is nice because it lends itself to the argument that this is mathematically tractable in that we are clearly specifying the mechanism/distribution. This is less clear in the fixation method where the empirically observed y_t follows no clear distribution. Finally, we should speak to the fact that we are omitting what appears to be "information gathering behavior". This was addressed in McMurray 2022. I will elaborate more in the discussion, but in short the idea that there is info gathering behavior information in the fixations violates the assumption that activation is running in parallel from visual stimuli. By introducing the saccade method, we are leaving the fixations as an entirely separate component with some potentially interesting avenues to pursue.

5 Simulations

Simulations were conducted to replicate the mechanics of a look combined with oculomotor delay, detailed in Figure ??. This section only address Target fixations with a four parameter logistic as detailed in Equation ??; simulations according to looks to competitors is treated in the appendix. We will begin by describing the process of simulating a single subject.

First, each subject randomly drew a set of parameters θ_i from an empirically determined distribution based on normal hearing participants in the VWP [?] to construct a subject specific generating curve, $f_{\theta_i}(t)$. It is according to this function that the decision to initiate a look at time t will subsequently direct itself to the Target with probability $f_{\theta_i}(t)$.

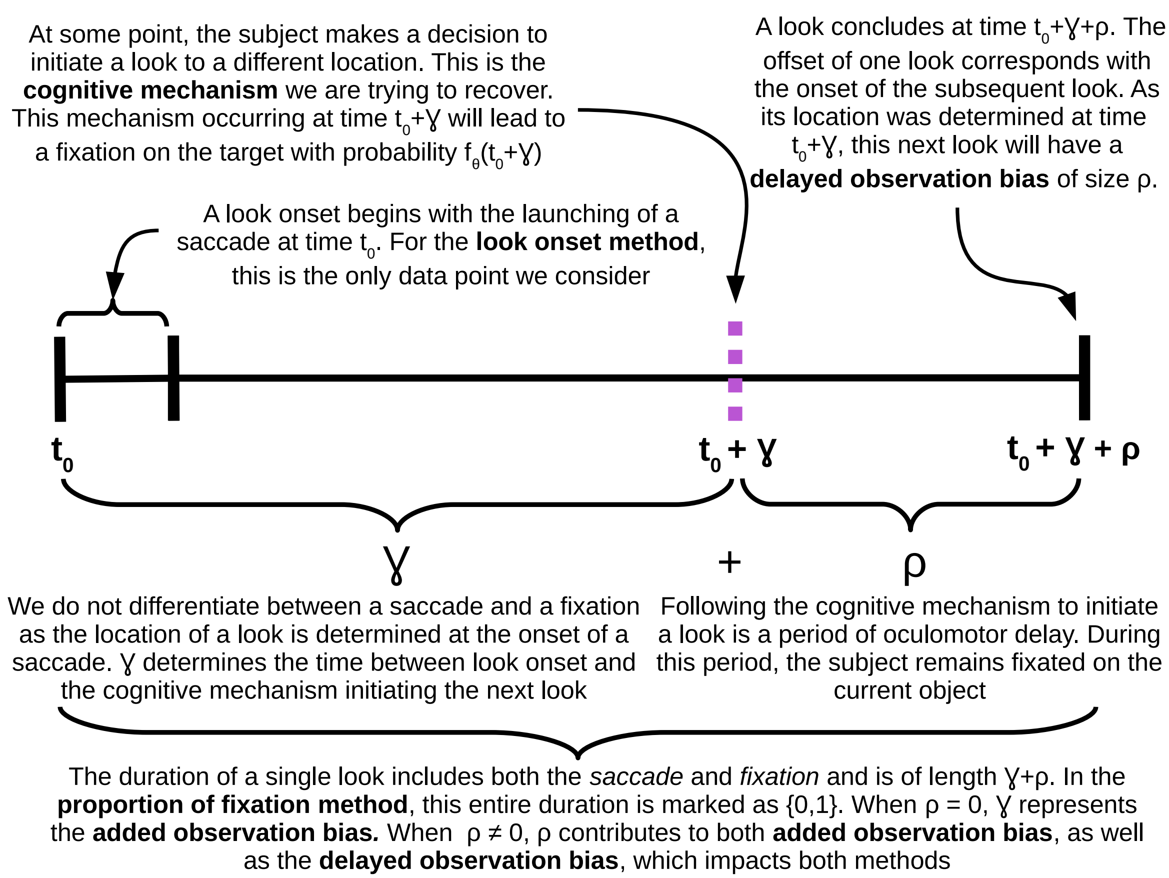
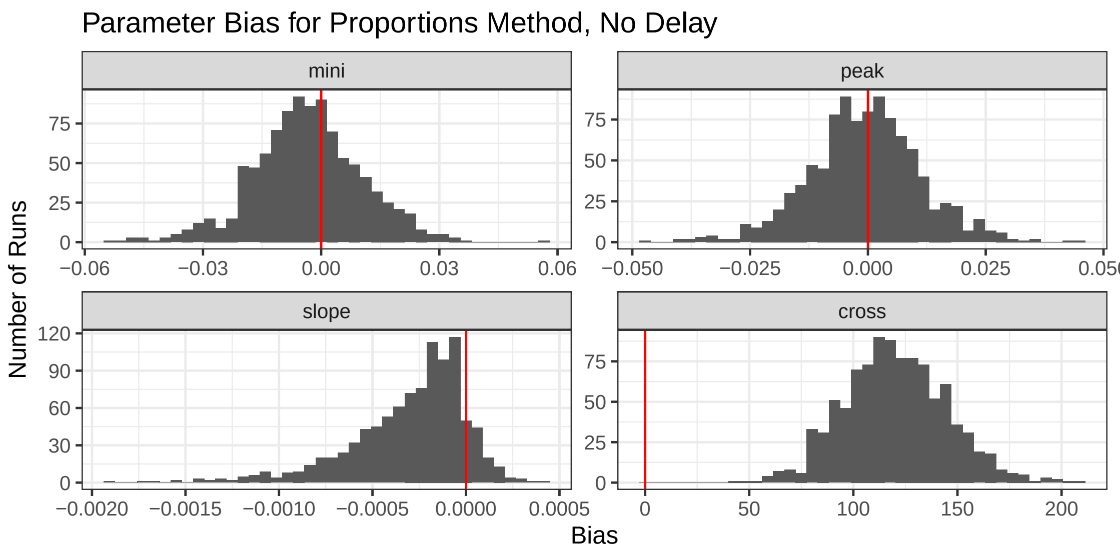


Figure 7: Anatomy of a look

5.1 No Delay

981 retained



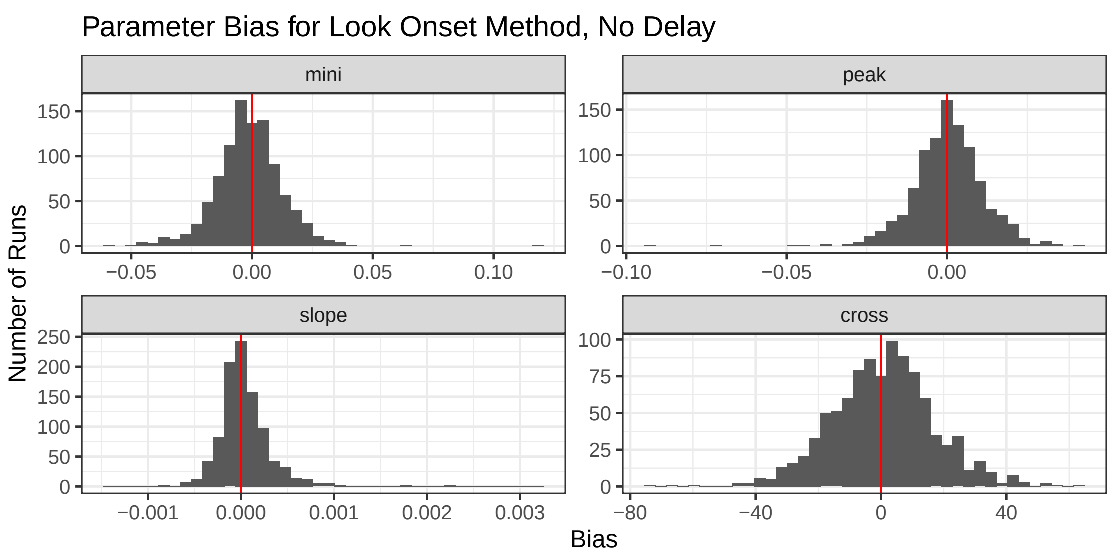


Figure 8: Parameter bias for no oculmotor delay. Part of me wants to flip these all so that look onset on top. These should also be as subfigures

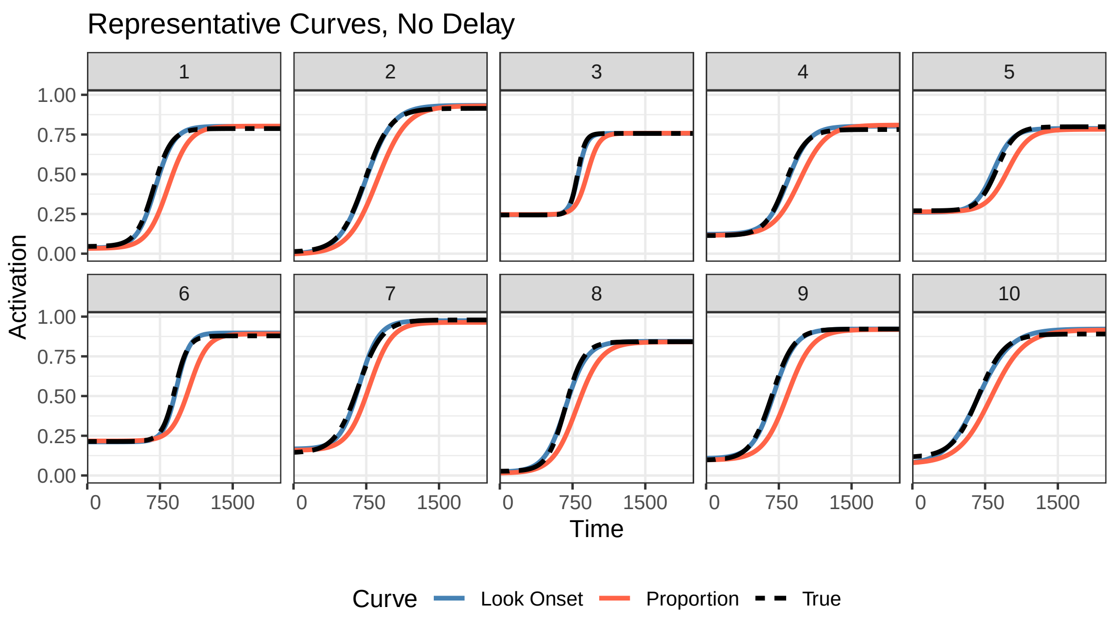


Figure 9: Representative curves for no oculmotor delay

5.2 Uniform Delay

973 retained

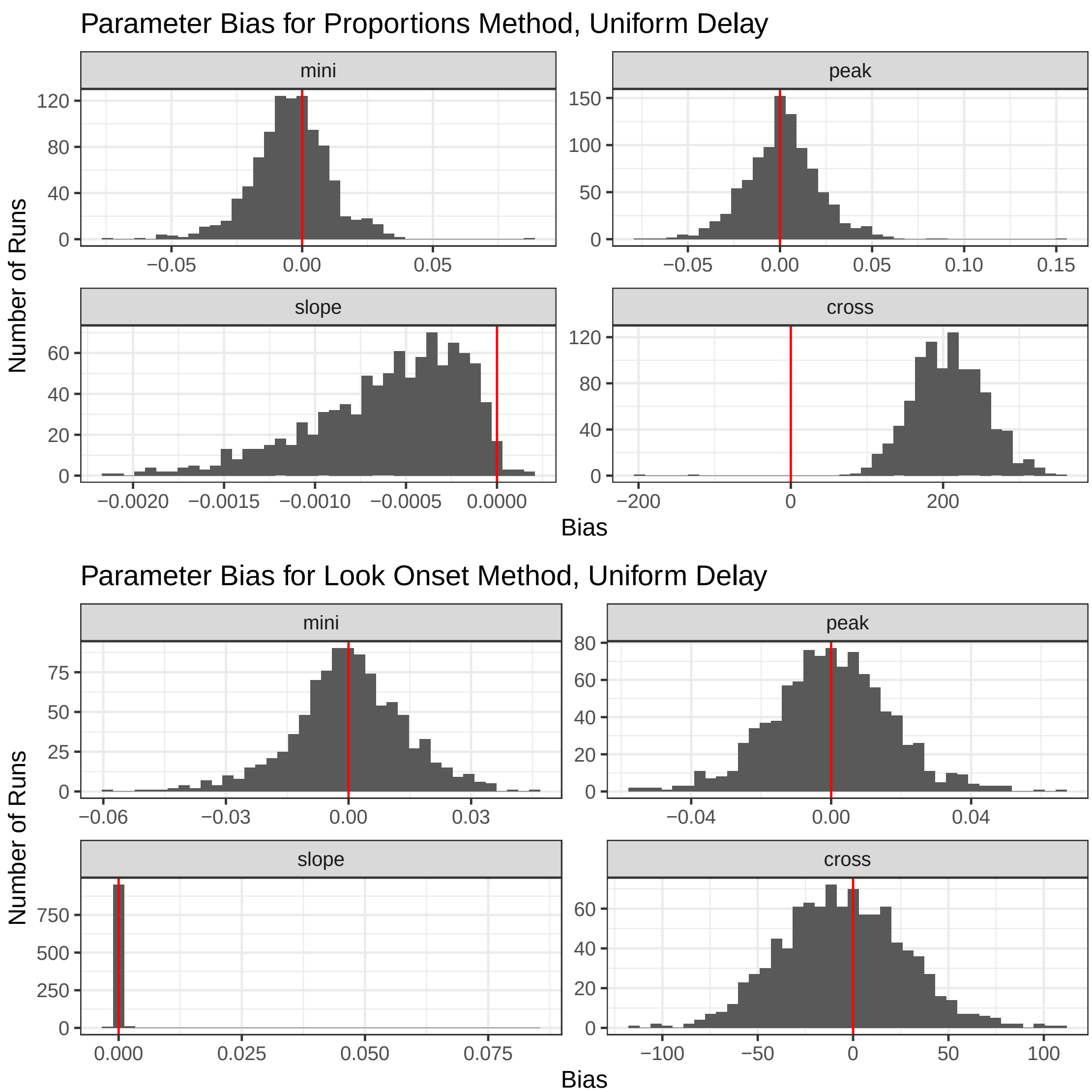


Figure 10: Parameter bias for uniform OM delay. These should be as subfigures

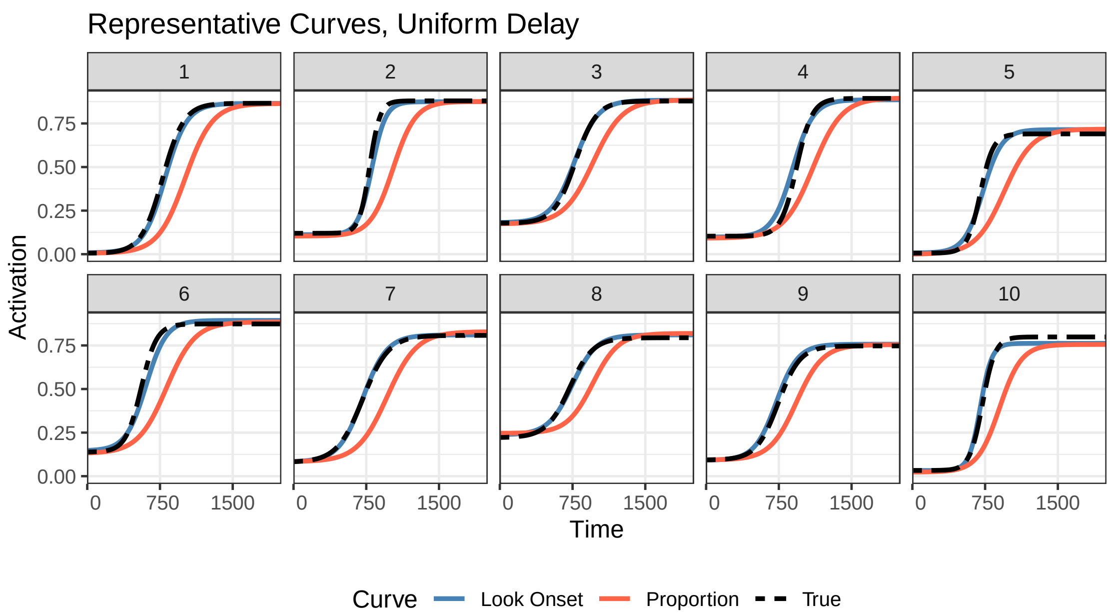


Figure 11: Representative curves for uniform oculmotor delay

5.3 Weibull Delay

981 retained

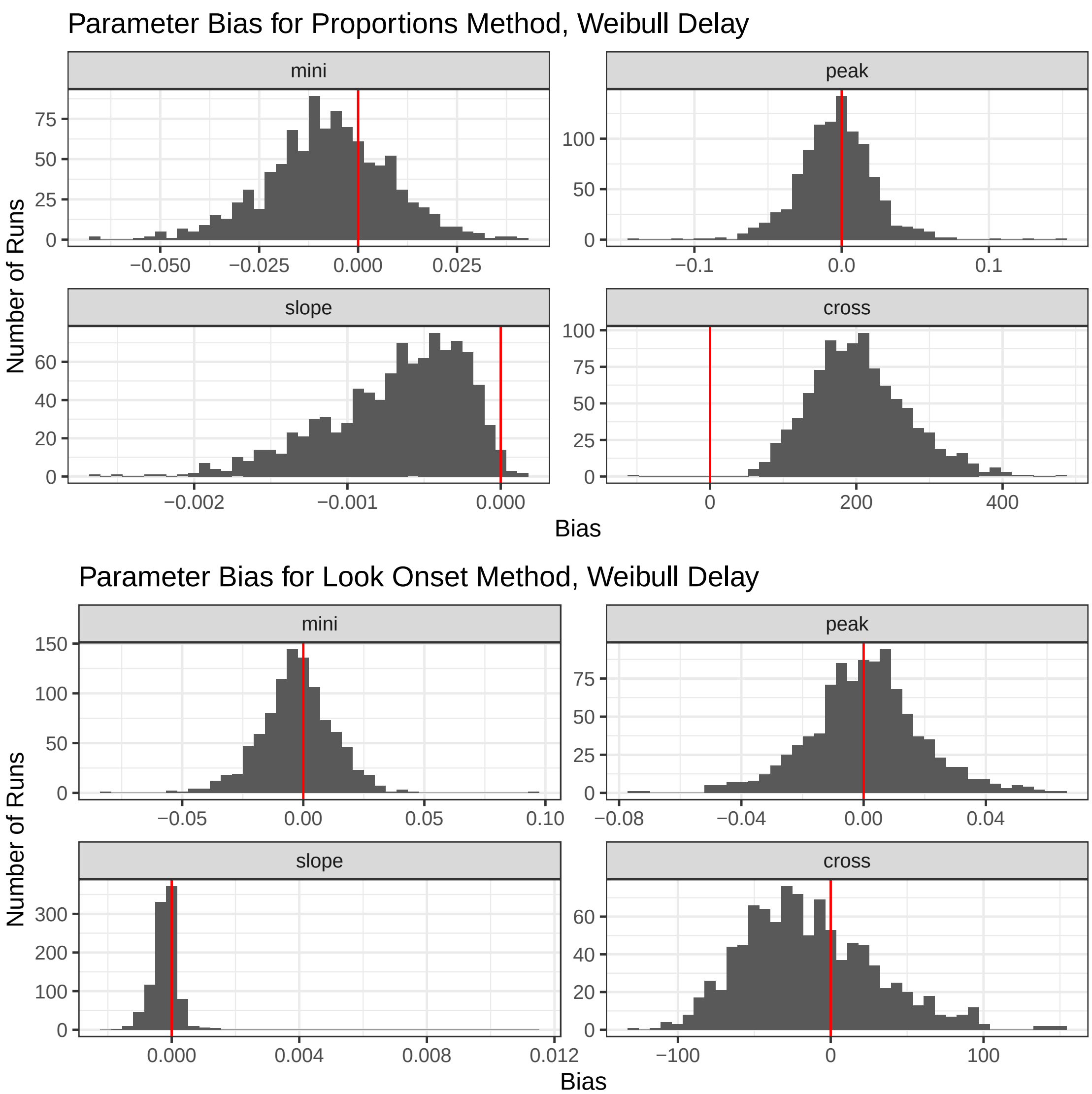


Figure 12: Parameter bias for weibull OM delay. These should be as subfigures

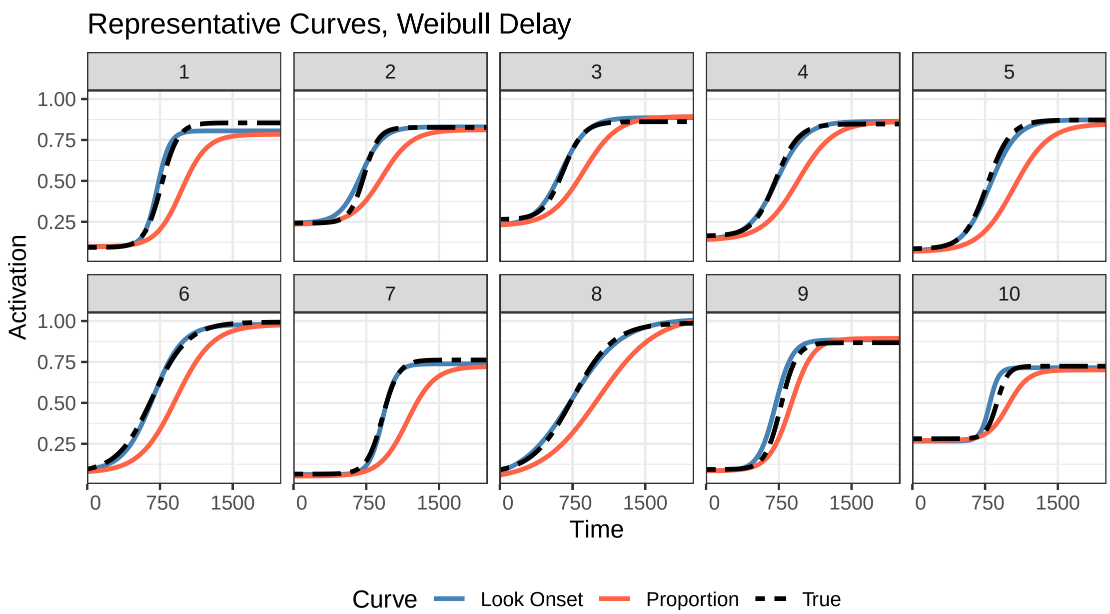


Figure 13: Representative curves for weibull oculmotor delay

5.4 Results

Curve	Delay	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Look Onset	No Delay	0.01	0.17	0.33	0.43	0.55	6.57
Look Onset	Uniform Delay	0.02	0.49	0.93	1.32	1.74	8.93
Look Onset	Weibull Delay	0.01	0.90	1.78	2.35	3.08	17.56
Proportion	No Delay	1.43	8.39	11.49	12.44	15.57	38.68
Proportion	Uniform Delay	3.35	20.70	29.48	31.61	40.09	126.54
Proportion	Weibull Delay	3.37	18.91	28.48	32.57	41.73	133.16

Table 1: Summary of MISE across simulations. I don't think I necessarily need (or want) all of those summary stats

5.5 Discussion

This section needs to be tightened and I have said some things elsewhere. Instead, let this be a general collection of thoughts for now.

I would like to speak a little bit more on the concept of "information gathering behavior". One of the primary benefits of the proportion method is that it indirectly captures the duration of fixations, with longer times being associated with stronger activation. This also becomes important when differentiating fixations associated with searching patterns (i.e., what images exist on screen?) against those associated with consideration (is this the image I've just heard?). There seems to be a general consensus also that longer fixations correspond to a stronger degree of activation, but a crucially overlooked aspect of this is the implicit assumption that fixation length and activation share a linear relationship. Specifically, insofar as the construction of the fixation curves is considered, a fixation persisting at 20ms after onset (and well within the refraction period) is considered identical to a fixation persisting at 400ms. More likely it seems this would be more of an exponential relationship, with longer fixations offering increasingly more evidence of lexical activation. By separating saccades and fixations at the mathematical level, we are able to construct far more nuanced models (one proposal, for example, might be weighting the saccades by the length of their subsequent fixation, or perhaps constructing a modified activation curve $f_{\theta(t)}(t)$ whereby the parameters themselves can accelerate based on previous information. But this is neither here nor there).

Speaking to the mathematical treatment, there is a wonderful simplicity in letting the saccades themselves follow a specific distribution, namely

$$s_t \sim Bin(f_\theta(t)) \tag{5}$$

or, with random oculomotor delay $\rho(t)$ (which I haven't really elaborated on as a separate mechanism),

$$s_t \sim Bin(f_\theta(t - \rho(t))) \tag{6}$$

This is in contrast to the fixation method, where the proportion of fixation curves can be described

$$y_t = \frac{1}{J} \sum z_{jt}. \tag{7}$$

Here, is there a clear distribution for what y_t follows? Under independence it may be the sum of binomials, but then what can be said about the relation of y_t to y_{t+1}, given that they may or may not share overlapping fixations from different trials? This is addressed to some degree in Oleson 2017, but this seems more of an ad hoc adjustment to account for this in retrospect. In contrast, the proposed saccade method makes no assumption of trial-level relationship and instead considers all saccades over all trials as binomial samples from the same generating curve in time.

This of course does ignore trial/word/speaker variability, but then perhaps it is time that we shift our language to speaking about a distribution of generating curves for a subject rather than a particular level of activation (note too that this utility is also reflected in the conversation regarding p-values against confidence intervals).

The arguments presented here has hoped to satisfy two goals, agnostic to the linking hypothesis or functions ultimately decided upon. Foremost is the recognition that saccades and fixations are governed by separate mechanisms, and treating them as such allows for fewer assumptions. For example, reconsider again the quote from Allopenna 1996:

"We made the general assumption that the probability of initiating an eye movement ot fixate on a target object o at time t is a direct function of the probability that o is the target given the speech input and where the probability of fixating o is determined by the activation level of its lexical entry relative to the activation of the other potential targets."

Under the saccade method, we omit the entirety of "and where the probability of fixating o is determined by the activation level of its lexical entry relative to the activation of the other potential targets" while still retaining the entirety of the utility in fitting *the same non-linear curves* to less of the data. This decoupling allows the typical time-course utility of the VWP to be used in conjunction with other methods treating aspects of the fixations separately.

Second to this, we have put a name to two important sources of potential bias in recovering generating curves in such a way as to be generalizable beyond the specifics of the assumptions of the simulation (both here and in McMurray 2022). The first, of course, addresses what was just discussed in the decoupling of saccade and fixation data. The utility of the second comes in that it makes no assumptions as to the source of the delayed observation, removing (possibly) unnecessary specifications between oculomotor delay and general mechanics when the goal is to simply recover the generating function. This may be less relevant when the goal of a study is to specifically address the mechanics of decision making (which itself seems to be difficult to pin down).

In short, what we have hoped to accomplish here is not to drastically change the original assumptions presented in Allopenna (1996) and elaborated upon in Magnuson (2019), but rather to qualify them in statistically sound ways. And really, that is pretty much it. Saccade method is neat, works the same way as the proportion of fixation method, has a more justifiable model while reducing assumptions and allowing room for others.

As a not really conclusion, I am sometimes left to wonder to what degree the proportion of fixation method was a "local minimum" is the pursuit of utilizing eye-tracking data. The proportion of fixations created an ostensible curve, prompting McMurray to establish theoretically grounded non-linear functions to model them. These, in turn, where shown to be suitable functions with which to model saccade data over a period of trials. Had saccades lent themselves so naturally to visualizing as the proportion of fixations, perhaps that is where we may have started.

6 Discussion

what have we learned?

Here are really the main takeaways.

1. We are all revisiting question of linking hypothesis
2. In the process of doing so, Bob identified some critical issues, revealing two distinct sources of bias
3. By introducing saccade method, we remove one source of bias and clearly delineate two separate but likely correlated mechanisms
4. This effectively keeps the assumptions from Allopenna and all of the benefits of constructing a function in time for activation, but also allowing room now for fixations to be used separately in a number of ways (length of fixation, latency to look, total fixations, etc.,)

7 limitations

probably good idea to keep running list of these all in one place

1. linking hypothesis/cognition curve
2. trace parameters maybe/general degrees of freedom
3. only evidenced on logistic, though for practical not theoretical reasons
4. adding parametric form (necessity for saccade method)
5. oculomotor delay, where to discuss

8 appendices

Here I am just including more or less random sections that either do not have a definite place yet in the main body of the paper, are part of what might be considered future work, or truly are things that belong in the appendix. Presented in no particular order

Appendix A

Treatment of empirical data from McMurray 2010 to get fixation and saccade curves, along with treatment of TRACE data

Appendix B

I'm not sure if appendix appropriate, but discussion on why double gauss/cohort not considered. This is primarily a consequence of failure to fit adequate models with `bdots`, arising from the fact that `gnls` is highly sensitive to starting parameters. I have demonstrated that they *can* be fit, but successful fits are able to be acquired with a huge range of parameters, bringing into question any validity. As the point of this paper is to demonstrate bias and counter saccade/fixation methods, this seemed an unnecessary addition.

Appendix C

Maybe catch-all for all things OM related. Originally included work showing that fixed delay simply results in horizontal shift, as well as investigation into how the amount of bias is a function both of the length of delay along with the derivative of the generating curve around the delay. Bias near the asymptotes has minimal impact relative to delay near the crossover point.

I think this would be interesting for future research but a bit beyond the scope to detail much here. Could expand on the idea if there was interest as I already have code written up that samples differentially at different time points.

Appendix E – Double Gauss simulations

Same results but for DG

8.1 No Delay

877 retained

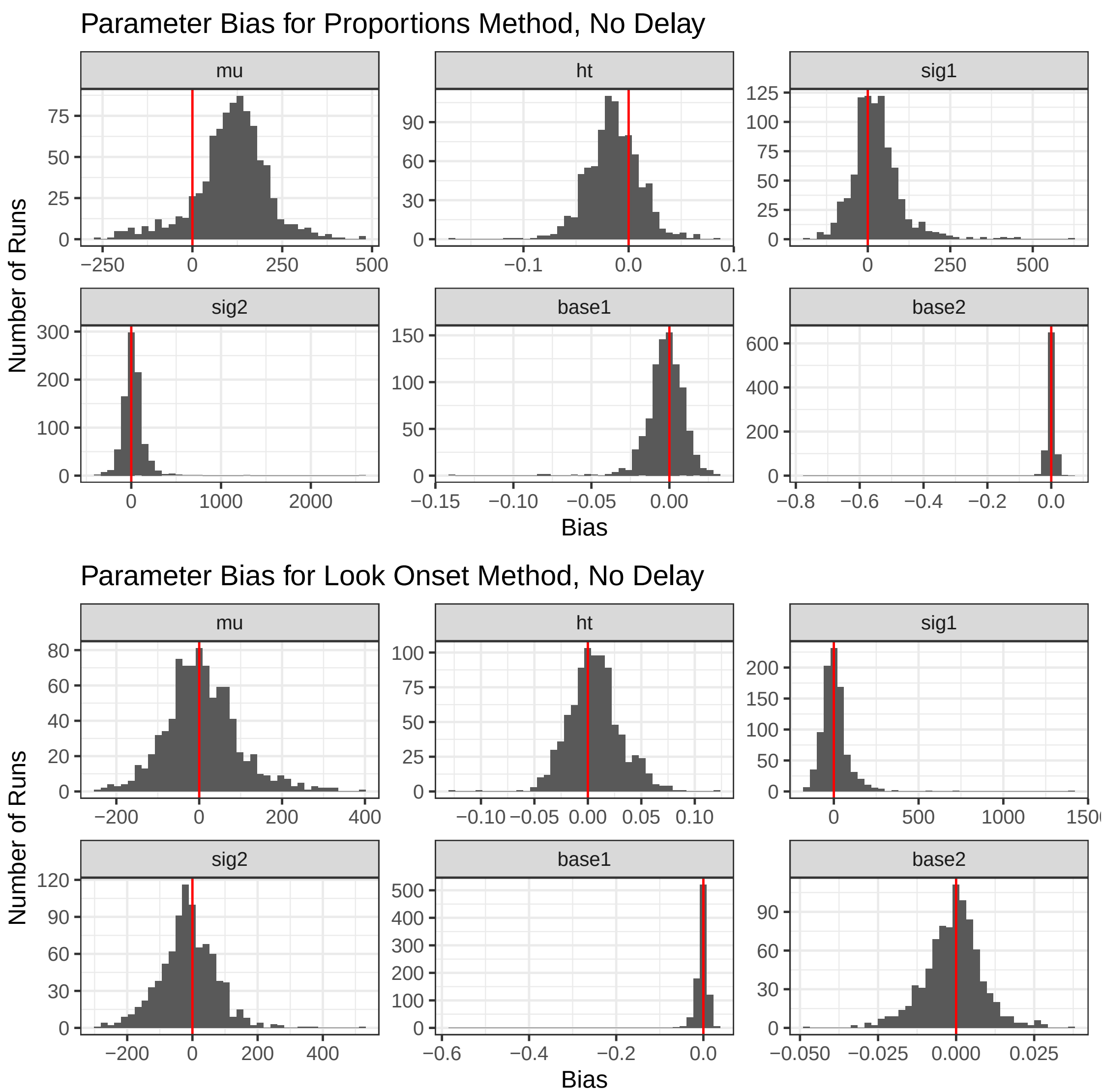


Figure 14: Parameter bias for no oculmotor delay

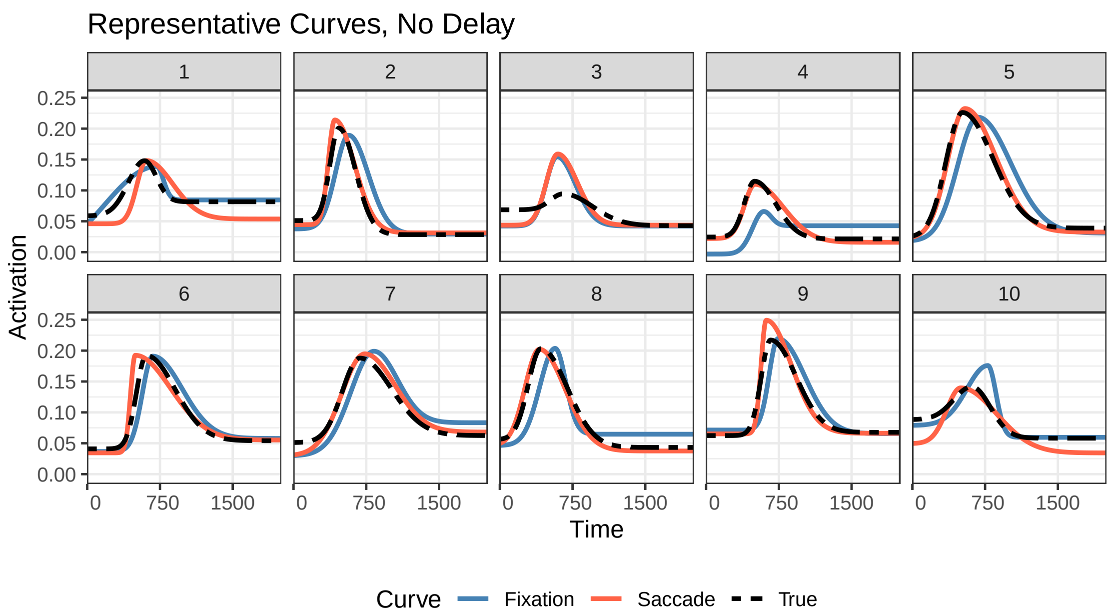


Figure 15: Representative curves for no oculmotor delay

8.2 Uniform Delay

844 retained

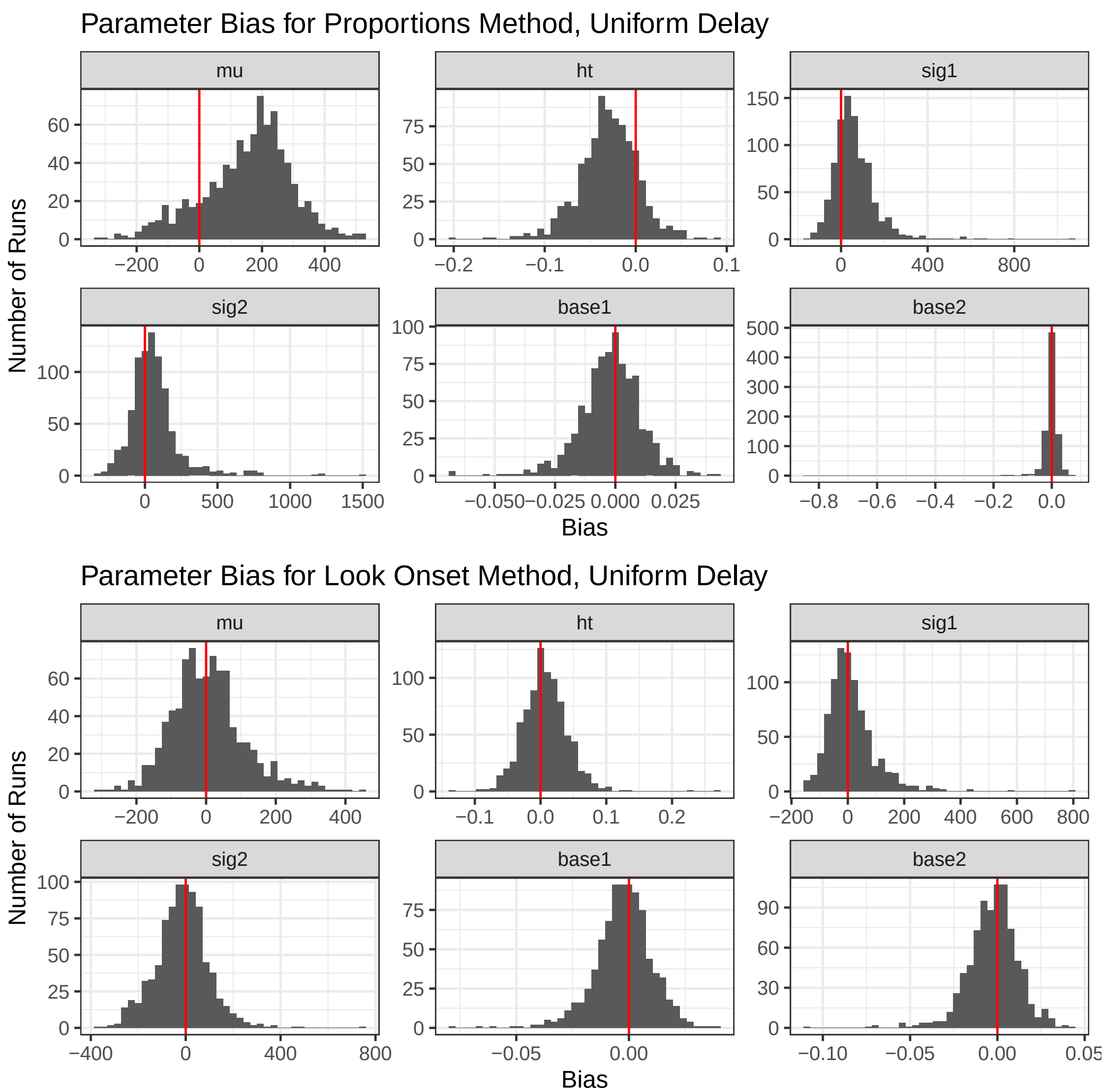


Figure 16: Parameter bias for uniform OM delay

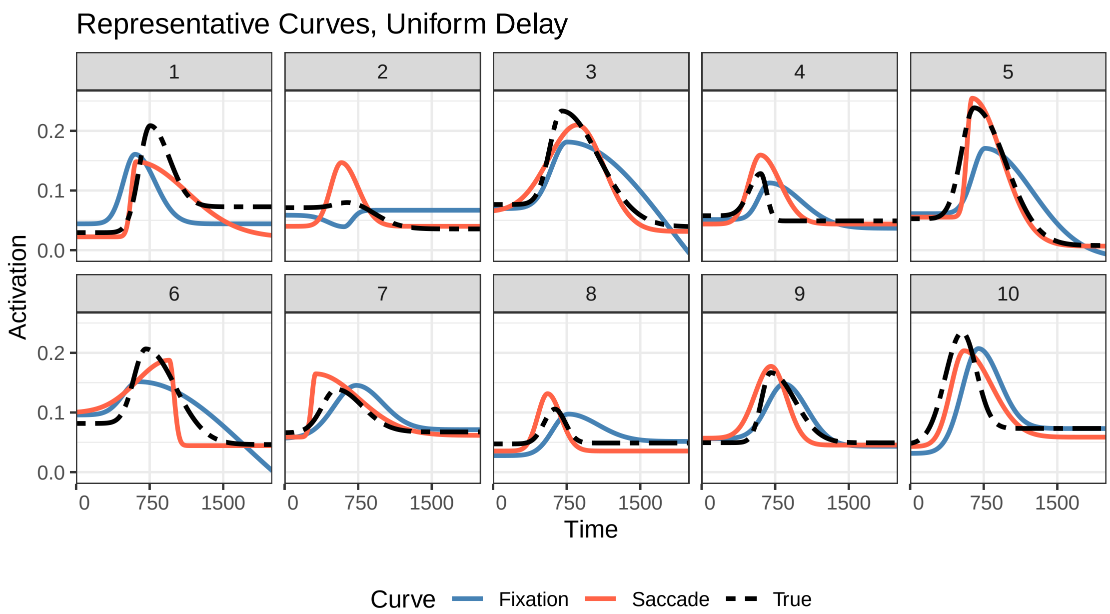


Figure 17: Representative curves for uniform oculmotor delay

8.3 Weibull Delay

861 retained

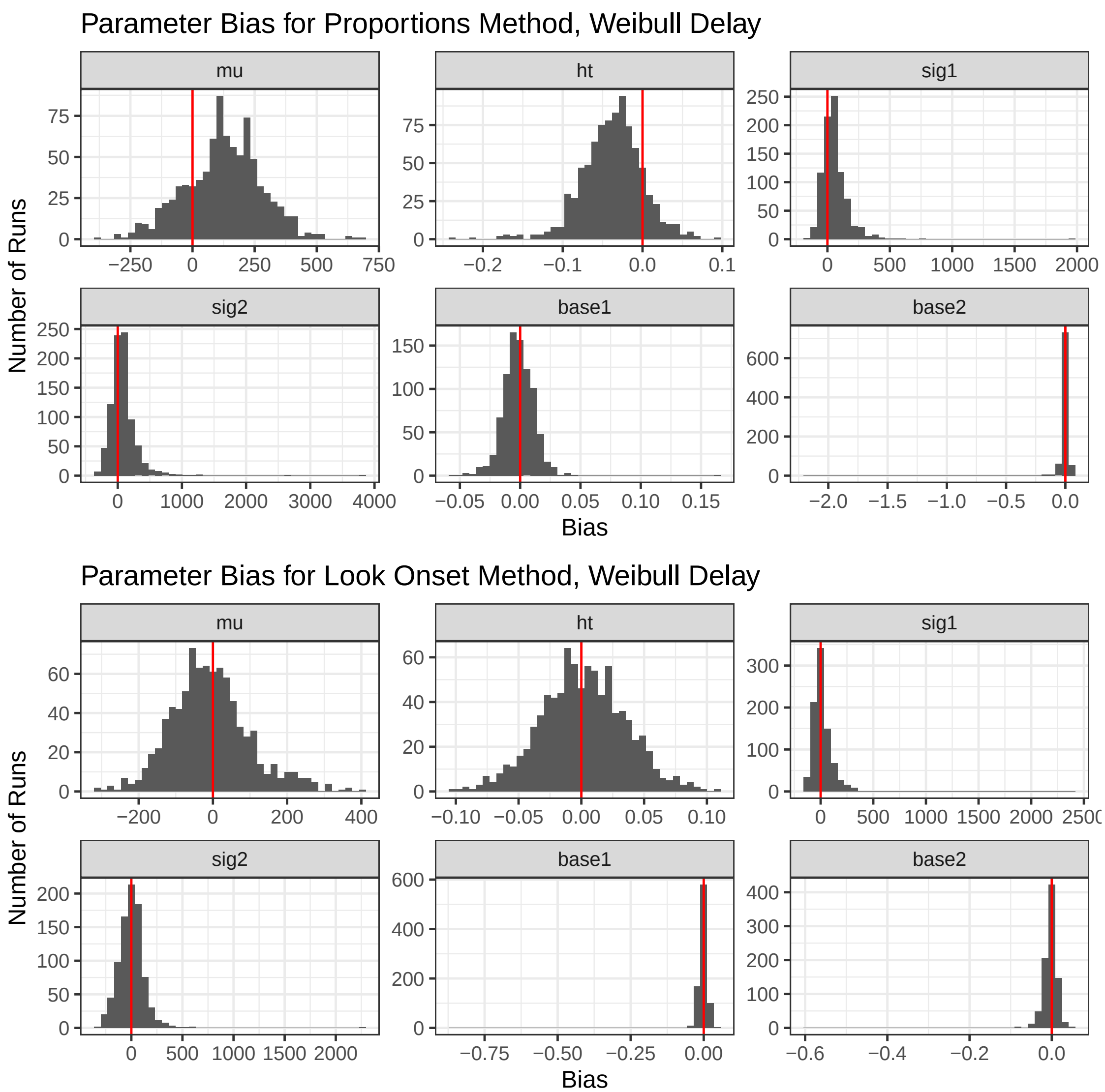


Figure 18: Parameter bias for weibull OM delay

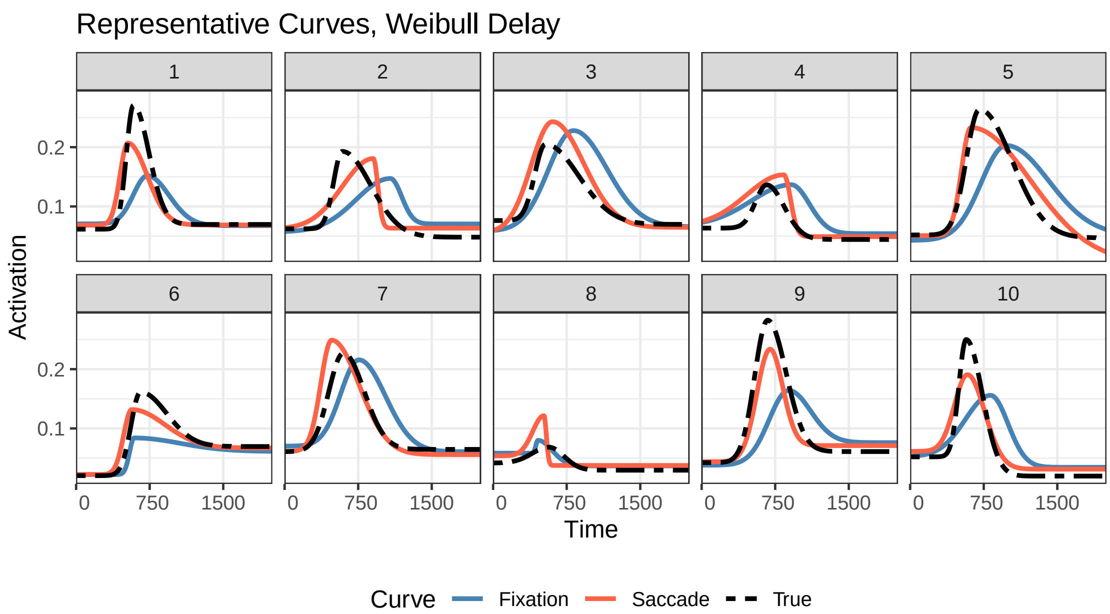


Figure 19: Representative curves for weibull oculmotor delay

8.4 Results

Curve	Delay	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Look Onset	No Delay	0.03	0.20	0.35	0.48	0.59	5.07
Look Onset	Uniform Delay	0.03	0.43	0.73	0.91	1.18	5.71
Look Onset	Weibull Delay	0.06	0.49	0.80	1.00	1.29	8.20
Proportion	No Delay	0.01	0.74	1.29	1.53	2.03	7.35
Proportion	Uniform Delay	0.05	1.27	2.25	2.73	3.69	12.26
Proportion	Weibull Delay	0.04	1.07	2.02	2.64	3.60	14.57

Table 2: Figures not as striking as logistic, but keep in mind that the scale of this is significantly smaller, peaking at around 0.15 and being close to zero elsewhere. R^2 is another metric available that

R^2 instead of MISE for logistic/doublegauss

,

The “minimum” for all of these is like -40 to -100, so mean is also not great for showing.

8.4.1 for logistic

Curve	Delay	1st Qu.	Median	3rd Qu.
Look Onset	No Delay	0.82	0.91	0.95
Look Onset	Uniform Delay	0.61	0.80	0.90
Look Onset	Weibull Delay	0.58	0.79	0.89
Proportion	No Delay	0.52	0.66	0.75
Proportion	Uniform Delay	0.11	0.40	0.58
Proportion	Weibull Delay	0.16	0.42	0.63

Table 3: R^2 for logistic, not as clear a hierarchy

8.4.2 for doublegauss

Curve	Delay	1st Qu.	Median	3rd Qu.
Look Onset	No Delay	0.82	0.91	0.95
Look Onset	Uniform Delay	0.61	0.80	0.90
Look Onset	Weibull Delay	0.58	0.79	0.89
Proportion	No Delay	0.52	0.66	0.75
Proportion	Uniform Delay	0.11	0.40	0.58
Proportion	Weibull Delay	0.16	0.42	0.63

Table 4: R^2 for double gauss