1 Where are we going?

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. Can elaborate on Magnuson 2019 to whatever degree relevant.

In particular, here we consider a contribution presented by McMurray 2022. From the abstract of this paper: "All theoretical and statistical approaches make the tacit assumption that the time course of fixations is closely related to the underlying activation in the system. However, given the serial nature of fixations and their long refractory period, it is unclear how closely the observed dynamics of the fixation curves are actually coupled to the underlying dynamics of activation."

This is a critical statement to have been made. Our intention is to revisit some of the questions raised in this survey and to start towards introducing a constructive path for moving forward. The assumptions made and the general arguments presented can be summarized briefly.

First, we begin with the assumption that there is some generating curve mediating the relationship between activation and saccade generations, and although mechanics are introduced to demonstrate increasingly complex behaviors, these themselves operate independently of the generating function. In this sense, the assumptions here are consistent with those presented by Allopenna 1996 in which word recognition runs parallel with input from visual stimuli. In other words, activation proceeds independently of what objects may have been seen or recognized, and having seen the target object at one instance has no accelerating effect on the rate of activation. Beyond this, there is an accounting for oculomotor delay using a fixed value of 200ms. Finally, there is the introduction of increasingly complex eye mechanics, differentiating in time the duration of the fixations and determining at what point in time the destination a particular saccade was made.

Being mostly narrative here, I won't elaborate too much further for now. But it suffices to address those points crucial for understanding the direction and purpose of the methodology being proposed. In short, the question that is being gotten at is this: in light of the assumptions just described and under increasingly complex conditions, are we able to recover the underlying dynamics of the system in question (activation) given that the "nature of the fixation record [is a] stochastic series of discrete and fairly long last physiologically constrained events?" Briefly, the answer is no.

McMurray notes that the typical, unspoken assumption implicit in VWP is what he calls the "high-frequency sampling" (HFS) assumption, which states that the underlying activation at some time determines the probability of fixation. This again parallels the assumptions made in Allopenna 1996: "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." McMurray goes on to note that this is "patently" untrue and is nothing more than a polite fiction.

Nonetheless, it is useful to compare the relationship of the underlying dynamics with the observed data in the context of the HFS assumption relative to other, more complex assumptions. Not sure how much detail is necessary here, but the critical things to note are this: the sources of bias introduced in the princess bride simulations can be described by two mechanisms: a fixed delay bias, through the introduction of oculomotor mechanics, and a random delay bias introduced by the random duration of fixations and their relationship to when the destination of a saccade movement was generated as well as when it was observed. Notably, the fixed delay bias resulted in no difficulty in recovering the generating curve under the HFS assumption; after accounting for a horizontal shift, the distribution of bias in the estimated generating parameters was generally symmetric and centered about zero. This was not the case when the duration of fixations had a direct relationship between the timing of the observed behavior.

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

1. The first we will call "delay observation bias". This can be either a random delay, as was implemented in the FBS/FBS+T methods, or a fixed delay, as was observed in all methods, but most notably under HFS and introduced via the imposed oculomotor delay
2. The second source of bias we call the "added observation bias". This involves the fact that we are "observing" data points, indicated with $\{0, 1\}$ at any time t without having observed any behavior associated with the generating curve at that time. This source of bias will be the primary emphasis for our proposal.

We consider first the delayed observation bias. In the simulations presented, this was generated through both a fixed delay, meant to simulate the effects of oculomotor delay, as well as a random delay period, introduced through a mechanism whereby once a fixation is "drawn", the subject remains fixed on a particular object for the full length of the fixation, with the following fixation's location determined at the *onset* of the previous fixation. This follows the idea that once a fixation is made, the subject begins immediately preparing to launch their next saccade.

McMurray demonstrated that under HFS with only a fixed delay, the generating curve was able to be recovered without bias. In reality, an oculomotor delay is either truly fixed, in which case recovery is trivial (and especially in the case of comparing generating curves between groups in which both the magnitude and location of observed differences will be preserved under a horizontal shift), or the delay has an aspect of randomness to it, in which case it simply adds to the already random delay that comes from uncertainty in knowing when the decision to launch a saccade is made. As such, we can evaluate the effect of the delayed observation bias by limiting ourselves to testing two cases: one in which there is a fixed delay (here assumed without loss of generality to be zero) and one in which the delay is random. This has the added benefit of freeing ourselves from having to account for any particular assumptions on the source of this delay, only to say that it exists.

The second source of bias introduced is what I call added observation bias and comes singularly from the fact that we do not differentiate between fixations and saccades in the observed data. To illustrate, consider a situation is which there is no delayed observation bias and that the probability that a saccade launched towards that target object at time t is directly determined by the activation of the target at time t, a la Allopenna. When we observe this saccade, s_t, we are directly sampling from the activation curve following some distribution at that point in time,

$$s_t \sim Bin(f_\theta(t)), \tag{1}$$

where $f_\theta(t)$ is assumed to be the activation curve (elaborated upon in a previous section, the "generating curve" in Bob's simulation). What, then, to make of the subsequent fixation at time $t+1$? Under the current method in which the proportion of fixations to the target are computed at each time (which we call the proportion of fixation method), we treat a saccade launched at time t identically with the subsequent fixation at time $t+1$, up 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 eye mechanics. An illustration of this bias is given in Figure 1

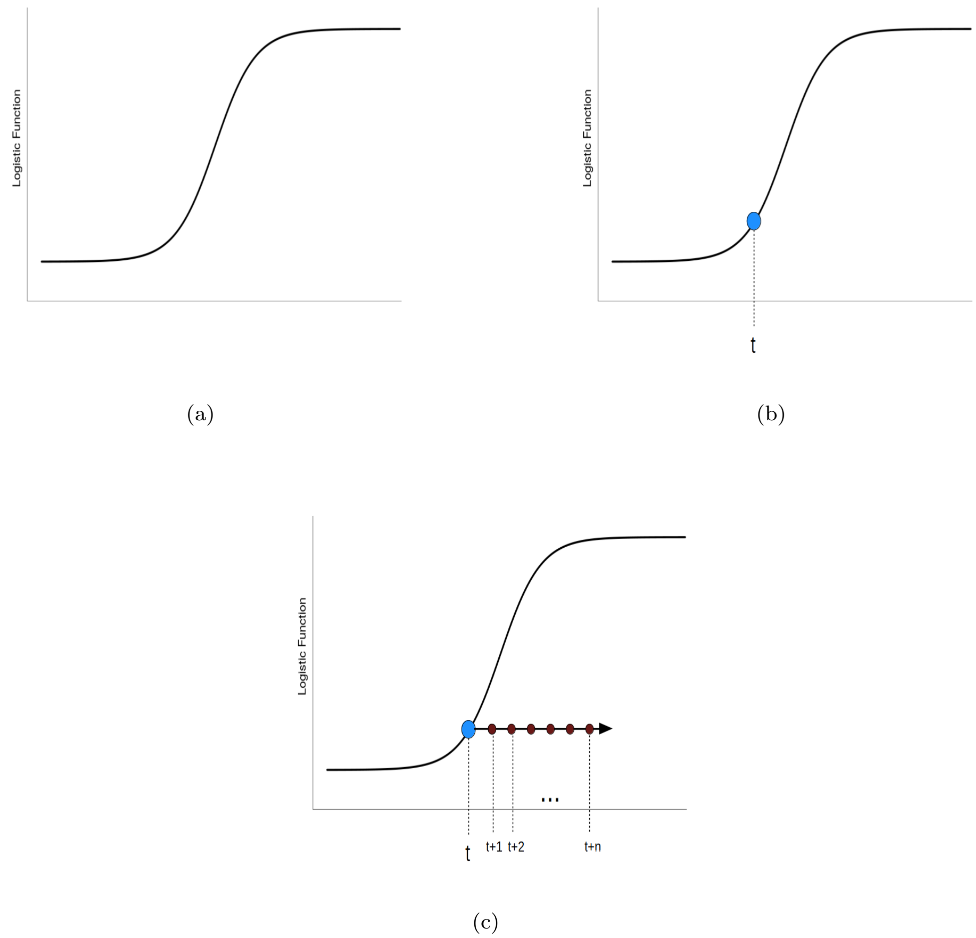


Figure 1: These illustrations can all be made larger (they were made for slides in an image editing program), but they illustrate the main point. **(a.)** here we see an example of a generating logistic function **(b.)** at some time, t, a saccade is launched (in the algorithm, a binomial is drawn with 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 we artificially inflate the *amount* of observed data. And in the particular case of the four parameter logistic function, we artificially *deflate* all of our observations. That is, as our function is monotone, 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.

While there is no immediate solution to the delayed observation bias, we argue that the added observation bias can be rectified by using *only* observations from saccades in the recovery of our generating curve. A few details on that next.

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 no 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 2. 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 and I think somewhere in Oleson 2017 ("Here fixations to each object can be considered an estimate of how strongly . 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.

2 Simulations

Here, we are going to attempt to isolate the two types of biased identified in the previous section, along with a comparison of the traditional proportion of fixation method with the proposed saccade method.

The first simulation will include no delayed observation bias – that is, a saccade launched at time t will be drawn directly from the generating curve at time t with probability $f_\theta(t)$. Here, $f_\theta(t)$ represents the generating function which we are ultimately hoping to recover. In this scenario, we should expect the saccade method to asymptotically provide an unbiased estimate of the generating curve. For the fixation method, any observed bias will be the direct consequence of the added observation bias.

In the second simulation, we will introduce a delay observation bias similar to that described in McMurray 2022. That is the duration of fixations will be random, following a gamma distribution with shape and scale parameters empirically determined (Farris-Trimble et al., 2014) (though in reality, any random distribution will do. Notably, the greater the skew the more pronounced the bias). Following a fixation, a saccade is generated, though with its probability of fixating on the target determined at the onset of the *previous* fixation. Here, both the saccade and fixation methods will demonstrate delayed observation bias, while the fixation method will continue to also demonstrate added observation bias.

These simulations differ from the original simulations presented in McMurray 2022 in a few regards. First, we have removed all together the oculomotor delay, instead keeping all of the delay observation bias random. This is a consequence of the trivial recovery that comes from horizontally shifting the underlying curve. Additionally, we have collapsed the complexity in generating eye movements into a single mechanic, consistent in that both are contributing to the same, random delay bias (that is, FBS and FBS+T differ in degree rather than kind). Finally, we limit our consideration to only a single generating function, the four-parameter logistic. This is for two reasons. First, we only wish to investigate the aforementioned sources of bias rather than any behavioral characteristics of the curves themselves. Second (and I can elaborate further) given the sensitivity of the fitting algorithm to starting conditions, the double-gauss remains more of a technical challenge to recover. This is a consequence of implementation rather than anything related to the theoretical discussion entertained here.

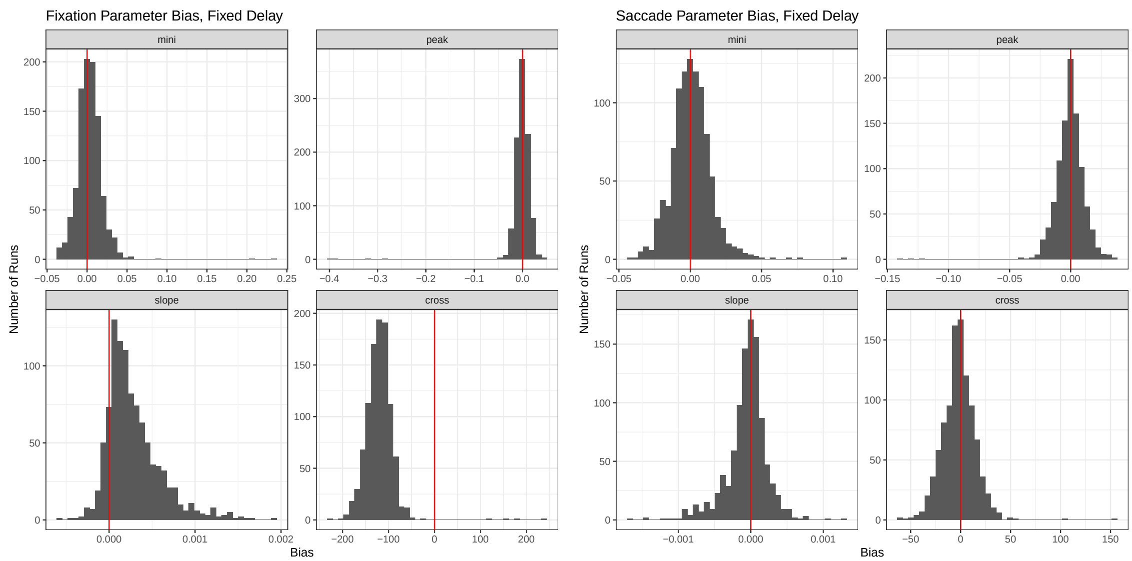
Similar to the original, an individual subject begins by drawing from an empirically determined set of parameters for their generating curves. Saccades were launched at random according to the details just outlined with probability determined by the generating curve. In each trial, a record was made of saccades launched, the time at which they launched, and where they had moved. Additionally, an indicator was computed every 4ms with either a 1 or a 0 to indicate if the current fixation was to the target or not, simulating the data generated through eye-tracking software. Fixations were repeated until the sum of fixations in a single trial exceeded 2000ms. Each subject performed 300 trials, and 1000 subjects were generated.

All saccade and fixation data was then fit to the four parameter logistic function with the R package `bdots` (v2) using the `logistic()` function. Given sensitivity to the starting parameters when fitting the curves and to ensure consistency in the fitting algorithm, both groups were given starting parameters `params = c(mini = 0, peak = 1, slope = 0.002, cross = 750)`. For curves fit to the fixation data, fitted functions with $R^2 < 0.8$ were discarded; for saccade data, fits were excluded if the base parameter estimate exceeded the base parameter or if the slope or crossover estimates were negative. Only subjects who passed both criteria were included. In all, 996 of the original 1000 subjects were kept under fixed delay conditions and 903 under the random delay conditions.

Lastly, each section will present a histogram of the observed bias in the recovery of the generating parameters, where the bias for each subject is computed as the generating parameter minus the recovered parameter. This means that positive bias is the result of underestimating the true parameter while negative bias is an overestimation. Each section will also present a representative collection of the fitted curves, both using the saccade and fixation methods, against the original, generating curves. Summary statistics on the quality of fits are reserved for the Results section where some more general comments are made.

2.1 Fixed Delay

Plots pretty self explanatory but I could elaborate here



(a) (b)

Figure 2: Distribution of parameter bias for fixation and saccade methods under fixed-delay simulation. The bias induced in the fixation method is all a consequence of the added observation bias./ We see evidence that added observation bias has the effect of "pulling" the curve at both ends, resulting in later crossover and less steep curves

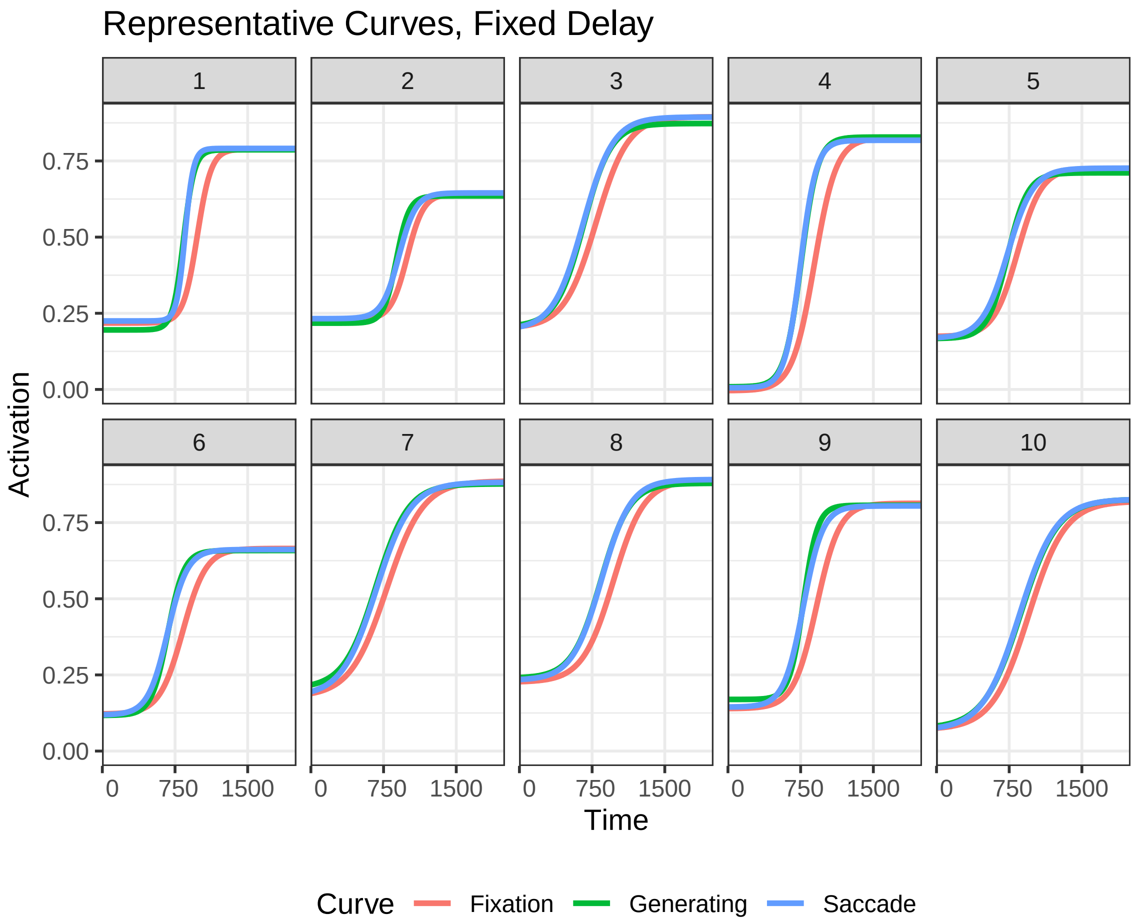
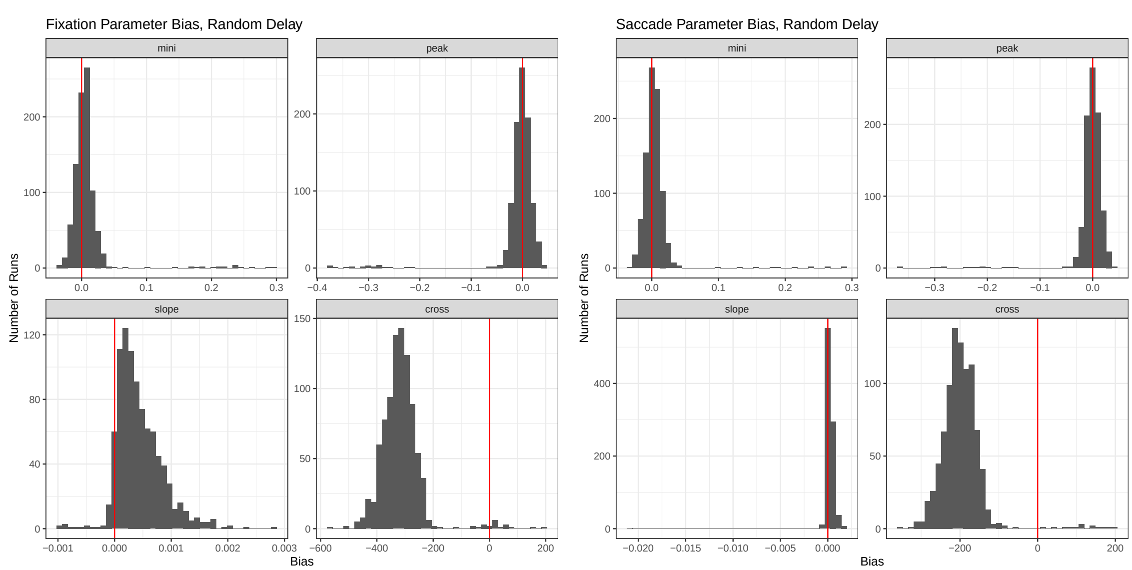


Figure 3: Representative collection of fixed-delay curve, including the generating function, as well as estimated curves from fitting data using fixation and saccade methods

2.2 Random Delay



(a) (b)

Figure 4: Distribution of parameter bias for fixation and saccade methods under random-delay simulation.

The bias induced in the fixation method is all a consequence of the added observation bias AND delay bias,

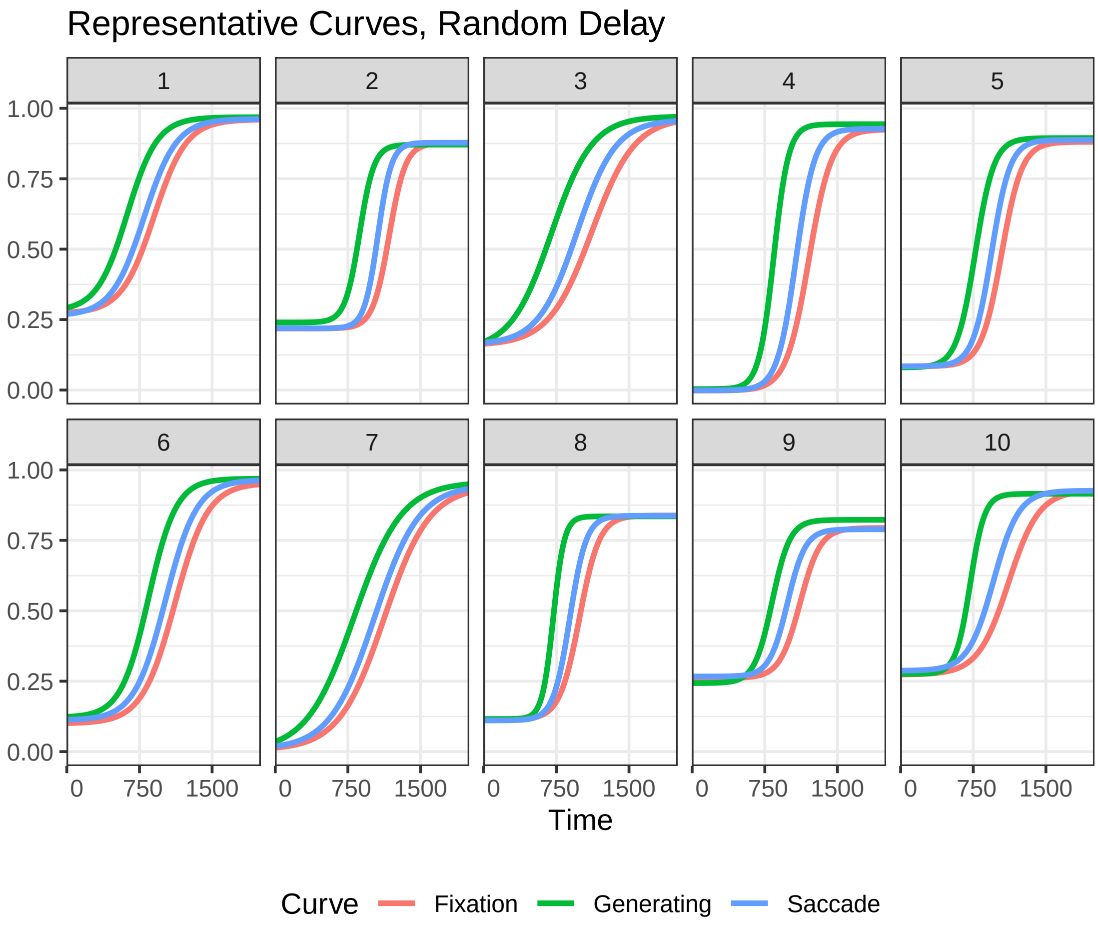


Figure 5: Representative collection of random-delay curve, including the generating function, as well as estimated curves from fitting data using fixation and saccade methods

2.3 Results

Perhaps unsurprisingly, Table 1 demonstrates that (1) Situations in which there is no delay between the generating function and observed behavior are easier to recover parameters and (2) the saccade method performed much better in all these cases. This table only includes MISE, I could add R^2, though the results will functionally be the same.

Curve	Delay	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Fixation	Fixed	1.95	8.18	11.40	13.28	15.98	215.67
Saccade	Fixed	0.01	0.16	0.32	0.52	0.56	78.22
Fixation	Random	20.25	50.95	68.60	73.08	90.92	192.56
Saccade	Random	5.74	21.42	29.29	33.40	40.63	185.79

Table 1: Summary of mean integrated squared error of the fits with their generating curves

2.4 Discussion

This section needs to be tightened but 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". There seems to be a general consensus 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 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).

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

$$s_t \sim Bin(f_\theta(t)) \tag{2}$$

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

$$s_t \sim Bin(f_\theta(t - \rho(t))) \tag{3}$$

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

$$y_t = \frac{1}{J} \sum z_{jt}. \tag{4}$$

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? 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).

Regardless of what linking hypothesis or functions are ultimately decided upon, the argument presented here has hoped to satisfy two goals. Foremost is the recognition that saccades and fixations are governed by separate mechanisms. Treating them as separate allows for fewer assumptions. 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. 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 (maybe) 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).

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..

3 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.,)

5. Showed that this was still consistent with continuous mapping models by agreement with TRACE