
Volume 123, Issue 2, May 2012
ISSN 0010-0277
pp. 199–334
ELSEVIER
COGNITION

Cognition 123 (2012) 325–334

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Discussion

Gaze step distributions reflect fixations and saccades: A comment on Stephen and Mirman (2010)

Richard S. Bogartz, Adrian Staub *

Department of Psychology, University of Massachusetts Amherst, Tobin Hall, Amherst, MA 01003, United States

ARTICLE INFO

Article history:
Received 22 June 2011
Revised 15 December 2011
Accepted 23 December 2011
Available online 14 January 2012

Keywords:
Eye movements
Distribution fitting

ABSTRACT

In three experimental tasks Stephen and Mirman (2010) measured gaze steps, the distance in pixels between gaze positions on successive samples from an eyetracker. They argued that the distribution of gaze steps is best fit by the lognormal distribution, and based on this analysis they concluded that interactive cognitive processes underlie eye movement control in these tasks. The present comment argues that the gaze step distribution is predictable based on the fact that the eyes alternate between a fixation state in which gaze is steady and a saccade state in which gaze position changes rapidly. By fitting a simple mixture model to Stephen and Mirman's gaze step data we reveal a fixation distribution and a saccade distribution. This mixture model captures the shape of the gaze step distribution in detail, unlike the lognormal model, and provides a better quantitative fit to the data. We conclude that the gaze step distribution does not directly suggest processing interaction, and we emphasize some important limits on the utility of fitting theoretical distributions to data.

1. Introduction

This article comments on Stephen and Mirman's (2010) approach to determining the nature of the cognitive processing that underlies eye movement behavior. The starting point of Stephen and Mirman's article is the observation that examining the shape of the distribution of a response variable can potentially provide insight into the nature of the processes that generated the values of that variable. In general, if the underlying processes take place in non-interactive, serially-ordered stages, the response variable will be distributed somewhere along a continuum containing the normal, exponential, and gamma distributions. A completely non-interactive, staged process will tend to result in an approximately normal distribution; if there is modest interaction between the outcomes of the stages, the distribution of the variable will move in the direction of the right-skewed (i.e., exponential and gamma) distributions. As an example of a serially-ordered process resulting in a normally-distributed response variable, consider the time it takes to type a sentence of a given length: This is simply the sum of the time it takes to type letter 1, letter 2, and so on. The central limit theorem tells us that this sum will tend to be distributed normally.

In contrast, there may be very strong interaction between processes or stages, such that the effect of each of several processes depends on the outcome of the others; we could model this by multiplying the values of the underlying variables. The product of random variables is generally not normal, but right-skewed. Indeed, this product distribution is sometimes approximated by a lognormal or power law distribution. As an example, consider the amount of money in an investor's portfolio after 10 years, without any new additions or withdrawals over this period. This final quantity will depend on the amount in the portfolio at the beginning of the period, and the rate of return each year. If each of these quantities varies

* Corresponding author. Address: Department of Psychology, University of Massachusetts, 430 Tobin Hall, Amherst, MA 01003, United States. Tel.: +1 413 545 5925; fax: +1 413 545 0996.

E-mail address: astaub@psych.umass.edu (A. Staub).

0010-0277/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2011.12.014

normally and independently, the distribution at the end of year 10 will have the characteristics of the lognormal and power law distributions, i.e., the modal value is at the extreme left, and there is an extremely long right tail.

The novel contribution of Stephen and Mirman (2010) was to use this distribution-based logic to address the nature of the cognitive processes that underlie eye movement control. They measured eye movements in three experimental tasks. Six subjects participated in a single feature search task, a conjunctive feature search task, and a variant of the visual world paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), in which subjects were presented with four objects on the screen and were instructed to click on the animal. The measure Stephen and Mirman analyzed is a fairly novel one in the eye movement literature: the *gaze step*, which is defined as the Euclidean distance, in pixels, between consecutive gaze positions sampled by an eye-tracking device (Aks & Sprott, 2003; Aks, Zelinsky, & Sprott, 2002; Stephen, Mirman, Magnuson, & Dixon, 2009). In the experiments, the tracker sampled at a rate of 500 Hz, i.e., every 2 ms. (With a constant sampling rate, gaze step is directly proportional to the eyes' angular velocity; e.g., when the gaze step over a 2 ms period is 10 pixels, the eyes are moving 10 times as fast as when the gaze step is 1 pixel.)

Stephen and Mirman (2010) claimed that if the cognitive processes underlying eye movement control are generally interactive, the observed distributions of gaze steps should be better fit by a power law or lognormal distribution than by an exponential or gamma distribution. Using a log likelihood goodness-of-fit criterion they found the lognormal to be the best-fitting distribution in 17 out of 18 cases (6 subjects $\times$ 3 tasks), with the power law distribution providing the best fit for the remaining case. They argued that "these findings reveal the interactive nature of cognitive processing and are consistent with theories that view cognition as an emergent property of processes that are broadly distributed over many scales of space and time rather than a componential assembly line" (p. 154).

We begin with an observation made by Stephen and Mirman (2010), who pointed out, based on visual inspection of these distributions on logarithmic axes, that there were "systematic departures of the empirical distributions from the best-fitting theoretical distributions" (p. 163). Stephen and Mirman refer to these departures as small, but what is noteworthy in their figures is that the form of the discrepancy between the observed data and the theoretical distributions is essentially the same for all 18 data sets, with all observed distributions showing an irregular two-hump shape that is not approximated by any of the candidate theoretical distributions (see Figs. 4–6 in Stephen & Mirman, 2010). Thus, while the distributions of gaze steps were closer to lognormal than they were to any of the other theoretical distributions explored in the paper, visual inspection reveals that the gaze step distributions do not, in fact, have the shape of lognormal distributions. Consistent systematic discrepancies between data and model ordinarily suggest that a search for an alternative model is in order. What kind of process could have generated the idiosyncratic shape that occurs for all 18 gaze step distributions?

Here we offer a simple account of the distribution of gaze steps, grounded in decades of empirical evidence about how the eyes move in tasks such as the ones under consideration here (see, e.g., Rayner, 1998; Rayner, 2009, for reviews). The eyes transition between two states: a fixation state, in which the gaze position is steady except for very small corrections (microsaccades; e.g., Martinez-Conde, Macknik, Troncoso, & Hubel, 2009) and noise introduced by tracker error; and a saccade state, in which the gaze position changes very rapidly as the eyes move to fixate a new location. Much more time is spent in the fixation state than in the saccade state, as each fixation tends to be much longer than the saccade that follows it. Fixations average in the range of 200–400 ms, with much variation due to the individual and the task; saccades average in the range of 20–50 ms, with much variation depending on the amplitude of the saccade. In addition, each saccade will involve an acceleration to a peak velocity, and then a deceleration. (See Gilchrist, 2011, for a review of the spatial and temporal characteristics of saccades.)

Our main argument is that a gaze step distribution is a single distribution in only a nominal sense: It is actually a mixture of two underlying distributions of gaze steps, one corresponding to fixations, the other to saccades. We proceed as follows. Making use of Stephen and Mirman's (2010) data,[1] we first illustrate the fact that gaze steps are organized temporally into long sequences of near-zero values (i.e., fixations), followed by small runs of larger values, which first increase and then decrease in magnitude (i.e., saccades). We then use a maximum likelihood technique to find the best-fitting mixture model for each of the gaze step distributions, making minimal assumptions about the nature of the two underlying distributions. We show that several positive results follow: the best-fitting parameters for the underlying distributions are in the range one would expect *a priori*, assuming that these underlying distributions correspond to fixations and saccades; the cumulative distribution functions (cdfs) of these mixtures closely mimic the shape of the observed cdfs but the best-fitting lognormal cdfs clearly do not; the mixture model always provides a better quantitative fit to the data than does the lognormal; and finally, the unexplained differences between tasks noted by Stephen and Mirman can be easily explained in terms of the parameters of the underlying mixtures.

2. Data

The reader is referred to Stephen and Mirman (2010) for details of the experimental method. Six participants completed 24 trials of each of three tasks: a single feature search task, a conjunction search task, and a variant of the visual world paradigm. Gaze positions were recorded by a remote SR EyeLink 1000 eye-tracker sampling at 500 Hz. The mean number of gaze steps (Euclidean distances between samples, in pixels) per participant-task combination was 36,590 (reflecting approximately 73 s of

[1] Thanks to Damian Stephen for providing us with these data.

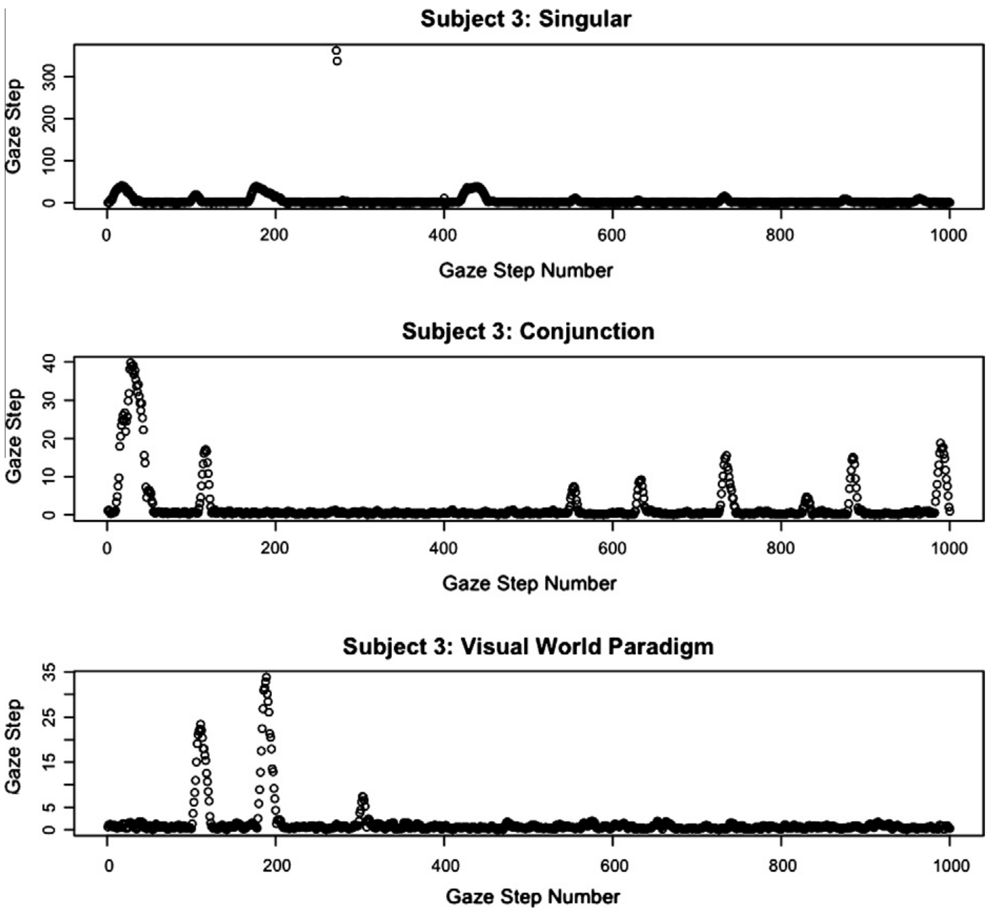

Fig. 1. The first 1000 gaze step values, as a function of the index of the observation, for one subject in the single feature search, conjunction search, and visual world tasks, in Stephen and Mirman's (2010) study.

eyetracking) with a standard deviation of 9703.24 and a range from 14,682 to 51,311.

3. Graphical display of gaze step sequences

Fig. 1 shows the first 1000 observations (2 s of data) for a single subject, in each of the three tasks. The index of the observation is on the *x*-axis, and gaze step in pixels is on the *y*-axis. The data are extremely well organized into long runs of near-zero values, followed by short periods in which there is a sudden rise and then fall in gaze step value. Clearly the eyes do indeed transition between two qualitatively distinct states, a long-duration state in which the eyes are relatively immobile, and a short-duration state in which the eyes move rapidly; within the latter state, the eyes rapidly accelerate to a maximum velocity (i.e., a maximum gaze step), and then rapidly decelerate.[2] Moreover, it is clear that a sequence of gaze steps will predict the next value with great precision; we note that across the 18 distributions, the serial correlation between each observation and the next ranged between .89 and .99. Note that this obvious temporal structure is unexplained on any modeling approach that focuses on the overall distribution of gaze steps, rather than on the sequence of events that gives rise to this distribution.

[2] It is worth pointing out that eyetracking devices determine when a saccade begins and ends by means of a change-in-velocity criterion; the beginning of a saccade is identified by determining the point at which velocity suddenly begins to increase, based on the distances between gaze positions across several successive samples.

4. Mixture modeling

We tested the assumption that the data for each of the 18 participant-task combinations could be modeled well as a mixture of two normal distributions, representing the fixations and saccades that are apparent in Fig. 1. For each of the 18 distributions of gaze steps, we estimated the mixture model parameters (the mean, μ_1 and μ_2, and variance, σ_1^2 and σ_2^2, of each of the two underlying normal distributions, and the proportion of observations drawn from the first of the two, π_1) by implementing the expectation–maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977), using the *normalmixEM()* function in the *mixtools* package (Benaglia, Chauveau, Hunter, & Young, 2009) for the R statistical programming language (R Development Core Team, 2009). We do not describe the EM algorithm in detail here; the interested reader is referred to the literature. To avoid biasing the parameter estimation process, we began each fit from the default parameters employed by the function, e.g., assuming that the mixing proportion of the two distributions is .5. (Note that this procedure does not guarantee that the optimal parameters are obtained; but as is evident below, an extremely good fit was obtained for every distribution.)

We note that prior to obtaining the mixture parameters for each subject-task combination, we eliminated gaze steps greater than 100 from each distribution. It was apparent that some of the obtained high values could only reflect tracker error or track loss; 100 pixels corresponds to well over 2° of visual angle, given the experimental parameters that Stephen and Mirman (2010) report. Thus, during the corresponding 2 ms the eyes would have to maintain a velocity of more than 1000°/s, which is substantially faster than the maximum saccade velocities reported in the literature (Boghen, Troost, Daroff, Dell'Osso, & Birkett, 1974; Yee et al., 1985). Stephen and Mirman apparently did not exclude any outliers, but we regarded this as necessary, given the apparent physiological impossibility of moving the eye at the rate some of the high values in the data would imply. The top panel of Fig. 1 contains two such points, and illustrates the fact that such extreme high values are not continuous with the previous or following gaze steps, consistent with the assumption that they represent track losses; instantaneous acceleration to a very high velocity is clearly not possible. Note, in any event, that our exclusion criteria were extremely conservative. The proportion of observations we eliminated was only .00082 overall, ranging from no observations to .00416 of the observations across the 18 distributions. (Moreover, as we demonstrate below in the section "Quantitative tests of model fit," the mixture model outperforms the lognormal model even without outlier exclusion.)

The obtained mixture parameters are shown in Table 1. We begin by noting several uniformities in these parameters. First, in all cases, μ_1 (the mean of the first distribution) is no more than .55 of a pixel. Thus, the EM algorithm always identified a component distribution consisting of very small values. Second, the proportion of observations that the EM algorithm identified as coming from this distribution was between .795 and .906. Third, the second distribution, which contributed between .1 and .2 of the observations, had a much larger mean value (ranging from 4.8 to 14.6 pixels), with a relatively large variance.

In sum, all fits converged on very similar parameter values. More importantly, these are values that would be expected based on the existing knowledge of fixation and saccade characteristics. Fixations account for approximately 80–90% of viewing time (and therefore, gaze steps), and there is very little eye movement during fixations; saccades account for the remaining viewing time, and eye movement is quite rapid during saccades. In addition, the eye must accelerate to its peak velocity (as shown in Fig. 1), and decelerate from it, so the gaze steps during saccades would be expected to be relatively variable, as velocity is variable.

5. Graphical display of fits

Next, we display cumulative distribution functions (cdfs) of the log of the observed gaze steps, in Figs. 2–4, for the single feature search, conjunction search, and visual world tasks, respectively. We superimpose on these cdfs the predicted cdfs generated from the best fitting mixture parameters in Table 1. In addition, we display predicted cdfs based on the best-fitting lognormal parameters, which we also obtained by maximum likelihood estimation.

Table 1
Best-fitting mixture model parameters for each distribution of gaze steps. Parameters μ_1 and μ_2 are the means of the two underlying normal distributions; σ_1^2 and σ_2^2 are their variances; and π_1 is the proportion of observations from the first distribution.

Feature search	μ_1	μ_2	σ_1^2	σ_2^2	π_1
Subject 1	0.209	4.844	0.019	63.506	0.903
Subject 2	0.550	7.492	0.144	57.082	0.861
Subject 3	0.517	9.417	0.097	68.239	0.861
Subject 4	0.343	6.788	0.045	63.667	0.906
Subject 5	0.436	11.169	0.074	222.863	0.881
Subject 6	0.360	9.349	0.052	157.109	0.851
Mean	0.403	8.177	0.072	105.411	0.877
Conj. search	μ_1	μ_2	σ_1^2	σ_2^2	π_1
Subject 1	0.211	6.203	0.019	50.979	0.868
Subject 2	0.506	8.216	0.119	95.620	0.858
Subject 3	0.461	9.513	0.080	73.148	0.853
Subject 4	0.344	6.680	0.045	48.623	0.900
Subject 5	0.449	10.098	0.079	175.995	0.853
Subject 6	0.344	8.757	0.045	150.783	0.859
Mean	0.386	8.245	0.065	99.191	0.865
Visual world	μ_1	μ_2	σ_1^2	σ_2^2	π_1
Subject 1	0.247	10.543	0.028	83.239	0.795
Subject 2	0.496	10.969	0.114	94.334	0.830
Subject 3	0.535	14.299	0.107	138.580	0.879
Subject 4	0.367	10.228	0.048	97.104	0.846
Subject 5	0.489	14.574	0.091	181.180	0.845
Subject 6	0.366	11.686	0.050	165.875	0.875
Mean	0.417	12.050	0.073	126.719	0.845

The most obvious feature of these graphs is that the mixture model does capture the details of the shape in every observed cdf, while the lognormal clearly does not. Moreover, the departure of the lognormal from the observed cdf is essentially identical in every case, as the lognormal cdf always has the same (almost linear) form. (Note that though the log–log plots in Stephen & Mirman, 2010, are graphically different from ours, a two-hump shape is evident in each of their plots, and the lognormal and power law pdfs do not capture this shape in their plots. Here we show cdfs of log gaze step because we regard these plots as more easily interpretable.) Clearly, the theoretically and empirically well-motivated assumption that a gaze step distribution is a mixture of two distributions is sufficient to capture the basic shape of the observed distributions, while the assumption that it is a single lognormal distribution does not capture this shape.[3]

We do note, however, that there are still systematic differences between the empirical cdfs and the cdfs generated

[3] The reader may note remarkable similarity between some of the plots. In part, this is due to the granularity in Stephen and Mirman's (2010) data; e.g. the five smallest values represented in each data file are 0, .1, .14142, .2, and .22361.

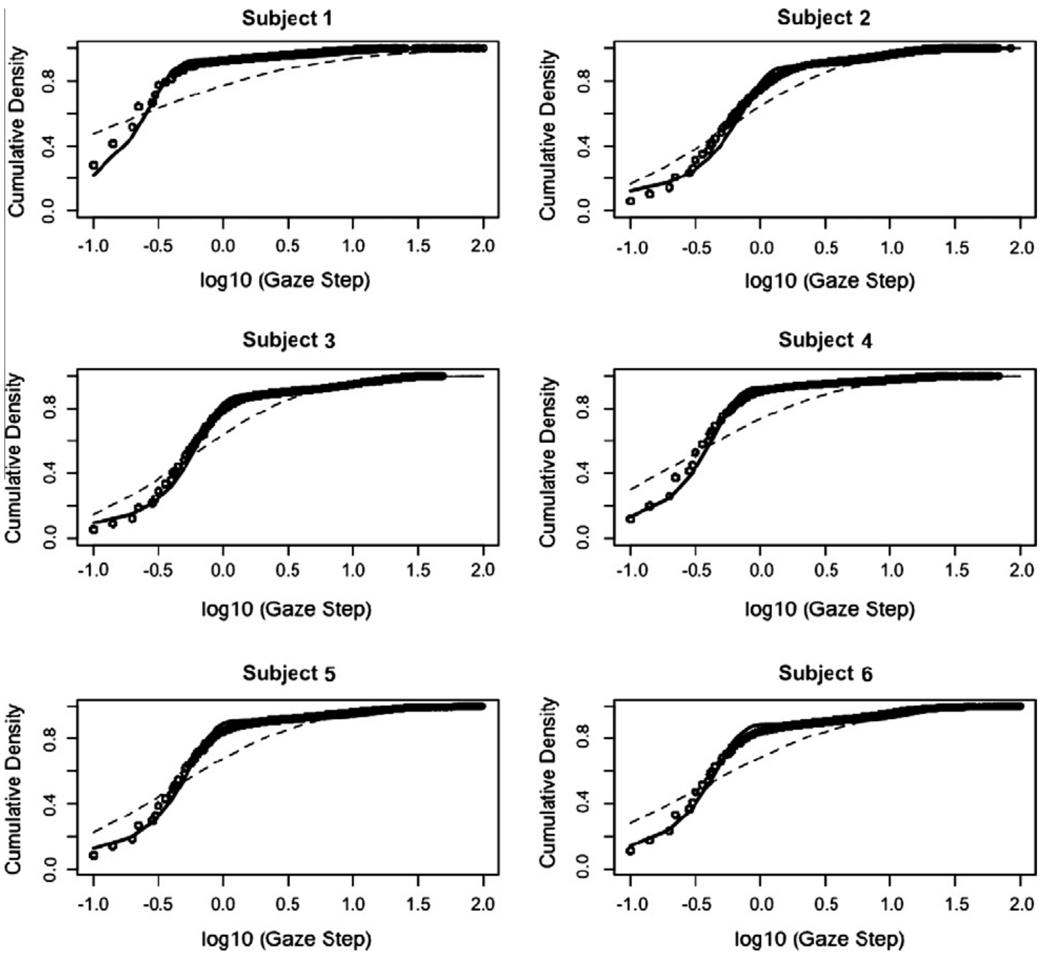

Fig. 2. Empirical cumulative distribution function (cdf) of the $\log_{10}$ of the gaze steps, for each subject in Stephen and Mirman's (2010) single feature search task. The data are represented by points; the cdf of the best-fitting mixture distribution is represented by a solid line, and the cdf of the best-fitting lognormal distribution is represented by a dashed line.

using the mixture model. These differences are very small compared to the deviation from the lognormal, but they are consistent with our assumption that in fact the two underlying distributions are probably not normal. Specifically, the gaze step distribution during fixations is bounded at 0, and the shape of the gaze step distribution during saccades will depend on the details of how the eyes accelerate and decelerate during saccades, and on the distribution of saccade amplitudes; we have abstracted away from these details here, for the sake of clarity and simplicity.

6. Quantitative tests of model fit

Though Figs. 2–4 show that the fit of the mixture model to the observed distributions is qualitatively much better than the fit of the lognormal model, it is conceivable that this is due only to the added flexibility the mixture model gains from having five free parameters, as opposed to two for the lognormal (see, e.g., Myung, 2000). To assess whether the fit of the mixture model is quantitatively better when it is penalized for its additional free parameters, we computed both the Akaike Information Criterion (AIC; Akaike, 1973) and Bayesian Information Criterion (BIC; Schwarz, 1978) statistics for each model fit. Each of these statistics begins with the log likelihood, and then penalizes the model based on the number of parameters. The penalty for increased flexibility (i.e., more free parameters) is generally greater with BIC than with AIC. For both statistics, a smaller value indicates a better penalized fit.

Table 2 presents the log likelihood, AIC and BIC statistics for each of the 18 distributions. The mixture model is the winner in all 18 cases, by both AIC and BIC. Indeed, the AIC and BIC statistics for the two models are never very close, despite the mixture model's penalty for extra parameters. Thus, it is safe to conclude that in all cases, the mixture model fits the data better not only qualitatively, but also quantitatively.

For completeness, in Table 3 we present the same information as in Table 2, but where no outlier values have been excluded. Again, the mixture model wins for all 18 distributions. The difference in fit is very slightly smaller than when outliers are excluded, which is expected, based on the fact that the lognormal model is better able to capture very high (and in this case, physiologically impossible; see Fig. 1) values.

7. Differences between tasks

We also address Stephen and Mirman's (2010) observation of differences in the gaze step distributions between

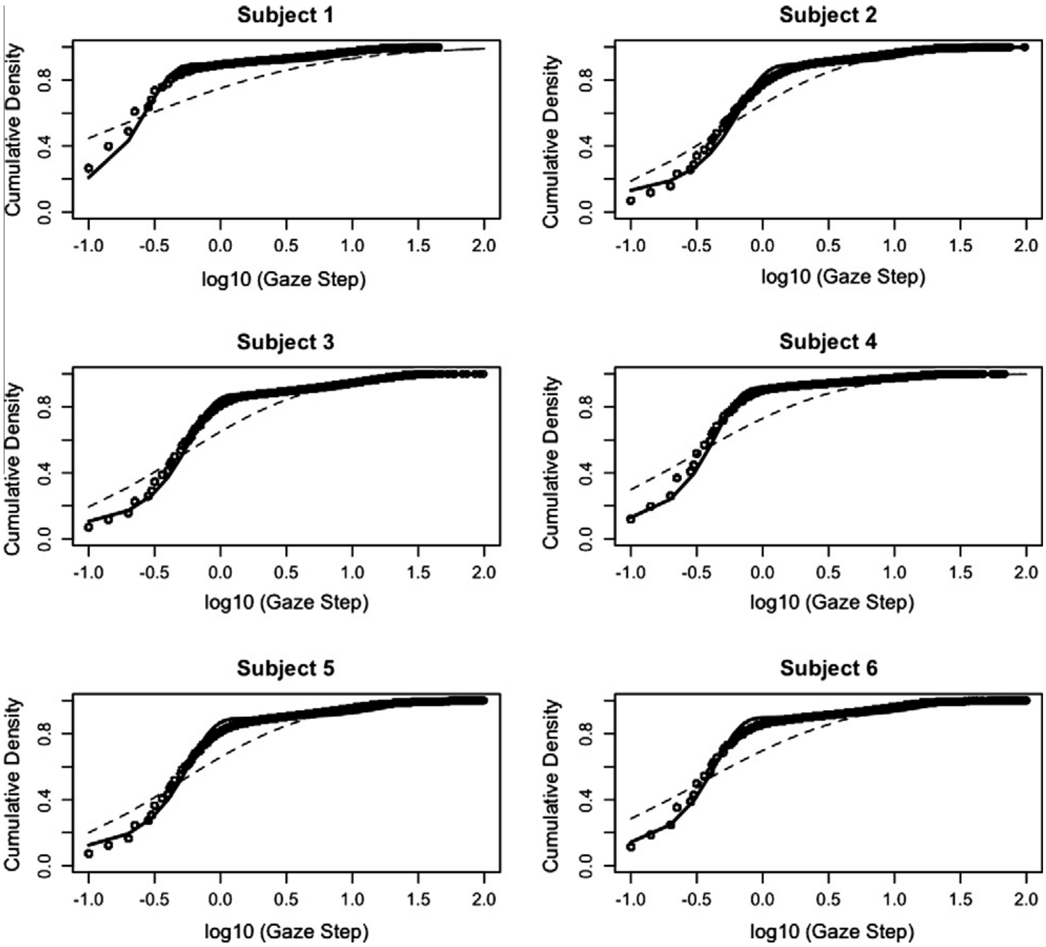

Fig. 3. Empirical cumulative distribution function (cdf) of the $\log_{10}$ of the gaze steps, for each subject in Stephen and Mirman's (2010) conjunction search task. The data are represented by points; the cdf of the best-fitting mixture distribution is represented by a solid line, and the cdf of the best-fitting lognormal distribution is represented by a dashed line.

tasks. They noted that while the advantage for lognormal over power law, in terms of maximum likelihood, was clear for the two search tasks, in the visual world task the fits were somewhat more equivocal between the two distributions. They did not address the question of why this might be the case, noting only that "task differences shifted cognitive dynamics along a continuum of interaction-dominance between lognormal and power-law distributions" (p. 162).

Examination of the best-fitting mixture parameters in Table 1, however, reveals differences between tasks that are unsurprising, given the nature of the tasks. We note that there are almost no differences between tasks in the mean of the fixation distribution (μ_1). This is as expected, based on the assumption that within-fixation eye movement is due entirely to oculomotor instability and to tracker error. On the other hand, the mean of the saccade distribution (μ_2) is substantially larger in the visual world task than in either of the search tasks. This, too, is consistent with our expectations. In Stephen and Mirman's (2010) visual world task, there were only four objects in the display, and these objects were relatively far apart, requiring some very long saccades. Because maximum saccade velocity increases with saccade amplitude (Boghen, Troost, Daroff, Dell'Osso, & Birkett, 1974; Gilchrist, 2011; Yee et al., 1985; see also Fig. 1), it is to be expected that some saccade gaze steps will be large when saccade amplitude is large.

The small differences between tasks in the proportion of samples drawn from the fixation distribution (π_1) are also consistent with this simple account. Because the visual world task elicits some long saccades, relatively more of the gaze steps will occur during saccades, and fewer will occur during fixations, i.e., π_1 will be relatively small. In general, it should be the case that when μ_2 is relatively large, π_1 will be relatively small, as both patterns arise when the task involves large-amplitude saccades. In sum, it is not only the case that the best-fitting mixture parameters are generally consistent with the assumption that there are two distributions corresponding to fixations and saccades; it is also the case that these parameters vary between tasks in a predictable way, based on expected differences in eye movement behavior in the different tasks.

8. Conclusion

We have demonstrated several phenomena relating to the gaze step variable. First, the existence of two eye movement states, fixations and saccades, is evident in plots of

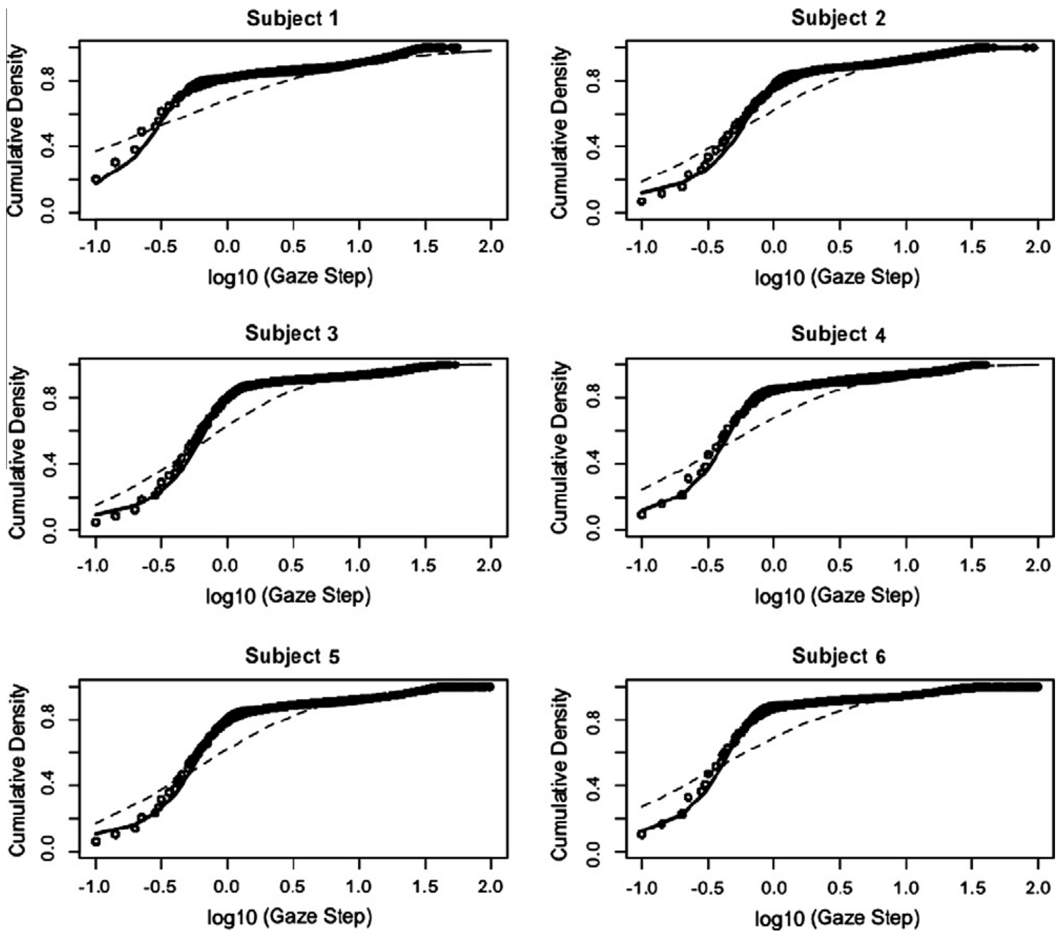

Fig. 4. Empirical cumulative distribution function (cdf) of the $\log_{10}$ of the gaze steps, for each subject in Stephen and Mirman's (2010) visual world paradigm task. The data are represented by points; the cdf of the best-fitting mixture distribution is represented by a solid line, and the cdf of the best-fitting lognormal distribution is represented by a dashed line.

the temporal sequence of gaze steps. Second, if a distribution of gaze steps is modeled as a mixture of two normal distributions, these two distributions have the characteristics one would expect from fixations and saccades. Third, for all 18 distributions considered here, the mixture model captures the idiosyncratic shape of the empirical distributions, while the lognormal model does not. Fourth, the mixture model provides a quantitatively better fit to all 18 distributions, even when it is penalized for its additional parameters. Finally, differences between tasks in the best-fitting mixture model parameters are consistent with basic differences between the tasks themselves.

At a broader level, we think that the important morals here are that researchers should carefully note qualitative departures from model fit, and should constrain their theorizing based on known facts about the processes underlying their data. The question of which quantitative model of some dependent variable provides the best fit to the data is of interest only as long as there are not aspects of the data that clearly suggest a failure of all the models under consideration. And the question of which statistical model provides the best fit should be considered against the background of the cognitive, neural, or motor processes that generated the data. In the present case, the nature of the gaze step distribution is mostly defined by basic facts about how the eyes move; it has its shape because the eyes are relatively stable for several hundred milliseconds, then rapidly, in only several tens of milliseconds, move to fixate a new location.

The observation that the distribution of gaze steps is a mixture, not a lognormal distribution, raises the question of how to regard Stephen and Mirman's (2010) substantive claim that interactions of cognitive variables underlie eye movement behavior. Clearly, we think that this conclusion is not supported by modeling of the distribution of gaze steps. However, this should not be taken to suggest that the claim of interactivity is wrong; we believe researchers will simply need to look elsewhere for evidence of such interactivity. There is a well-established modeling tradition in eye movement research (e.g., in reading: Engbert, Nuthmann, Richter, & Kliegl, 2005; Reichle, Rayner, & Pollatsek, 2003; in scene perception: Itti & Koch, 2001; Nuthmann, Smith, Engbert, & Henderson, 2010), and these models, which differ from each other in many details, treat fixation duration and saccade targeting as the phenomena in need of cognitive explanation. With respect to these variables, the question of interaction or non-interaction can be meaningfully posed. Are the durations of eye fixations and the targeting of saccades influenced by interactions between cognitive processes? The fact that the

Table 2
Log likelihood, AIC, and BIC statistics for the best-fitting mixture model and for the lognormal model, for each distribution of gaze steps.

Feature search	Mixture model			Lognormal model		
	loglik	AIC	BIC	loglik	AIC	BIC
Subject 1	−2871	5753	5765	−8305	16,613	16,618
Subject 2	−51,363	102,736	102,779	−56,695	113,394	113,412
Subject 3	−45,789	91,589	91,632	−57,208	114,421	114,438
Subject 4	−20,618	41,245	41,289	−36,156	72,316	72,333
Subject 5	−24,820	49,650	49,691	−31,625	63,254	63,270
Subject 6	−39,780	79,570	79,614	−49,922	99,848	99,866
Mean	−30,874	61,757	61,795	−39,985	79,974	79,989
Conj. search	loglik	AIC	BIC	loglik	AIC	BIC
Subject 1	−9893	19,795	19,808	−16,880	33,763	33,768
Subject 2	−55,777	111,564	111,608	−61,572	123,148	123,166
Subject 3	−47,966	95,941	95,985	−61,370	122,743	122,761
Subject 4	−23,651	47,311	47,355	−41,614	83,232	83,249
Subject 5	−43,736	87,481	87,524	−51,130	102,264	102,281
Subject 6	−41,623	83,255	83,300	−54,712	109,429	109,446
Mean	−37,107	74,225	74,263	−47,880	95,763	95,779
Visual world	loglik	AIC	BIC	loglik	AIC	BIC
Subject 1	−25,049	50,109	50,121	−30,335	60,673	60,678
Subject 2	−41,309	82,628	82,670	−46,674	93,352	93,369
Subject 3	−32,794	65,599	65,640	−42,131	84,266	84,283
Subject 4	−12,954	25,919	25,957	−17,472	34,949	34,964
Subject 5	−34,675	69,360	69,401	−42,186	84,377	84,393
Subject 6	−30,672	61,354	61,397	−44,447	88,898	88,915
Mean	−29,576	59,161	59,198	−37,208	74,419	74,434

Table 3
Log likelihood, AIC, and BIC statistics for the best-fitting mixture model and for the lognormal model, for each distribution of gaze steps, without outlier exclusion.

Feature search	Mixture model			Lognormal model		
	loglik	AIC	BIC	loglik	AIC	BIC
Subject 1	−6506	13,021	13,062	−8975	17,953	17,969
Subject 2	−37,995	76,000	76,042	−40,536	81,076	81,092
Subject 3	−47,239	94,489	94,532	−57,250	114,505	114,522
Subject 4	−14,994	29,998	30,039	−25,723	51,450	51,466
Subject 5	−25,604	51,218	51,259	−31,764	63,532	63,549
Subject 6	−48,284	96,577	96,620	−51,965	103,934	103,951
Mean	−30,104	60,217	60,259	−36,035	72,075	72,092
Conj. search						
Subject 1	−11,038	22,086	22,128	−16,930	33,865	33,881
Subject 2	−57,175	114,361	114,405	−61,702	123,409	123,426
Subject 3	−50,073	100,156	100,200	−61,467	122,938	122,955
Subject 4	−18,075	36,161	36,202	−27,292	54,589	54,606
Subject 5	−30,574	61,159	61,200	−34,454	68,912	68,928
Subject 6	−48,729	97,468	97,512	−55,914	111,832	111,849
Mean	−35,944	71,898	71,941	−42,960	85,924	85,941
Visual world						
Subject 1	−25,310	50,631	50,672	−30,345	60,694	60,710
Subject 2	−41,994	83,997	84,039	−46,688	93,379	93,396
Subject 3	−32,794	65,599	65,640	−42,131	84,266	84,283
Subject 4	−12,954	25,919	25,957	−17,472	34,949	34,964
Subject 5	−33,067	66,144	66,185	−39,686	79,375	79,392
Subject 6	−36,155	72,320	72,363	−45,573	91,151	91,168
Mean	−30,379	60,768	60,809	−36,982	73,969	73,985

distribution of gaze steps is a mixture of a fixation distribution and a saccade distribution is entirely consistent with a positive answer to this question. The eyes alternate between a fixation state and a saccade state, but it is entirely possible that, e.g., the time that the eyes spend in the fixation state is determined by complex interactions between cognitive processes.[4]

One notable empirical observation, in this regard, is that distributions of fixation durations in reading (Staub, White, Drieghe, Hollway, & Rayner, 2010; White & Staub, 2011) and in the visual world paradigm (Staub, Abbott, & Bogartz, submitted for publication) have a single peak and show only a modest amount of right skew, and are well fit by the ex-Gaussian distribution (Ratcliff, 1979). Indeed, Staub et al. (2010) and White and Staub (2011) have pointed out that fixation duration distributions in reading are substantially less skewed than are response time distributions in many behavioral paradigms. Thus, application of Stephen and Mirman's (2010) distribution-based logic would suggest that these durations are not generated by broadly interactive processes. We do not wish to endorse non-interactivity on this basis, however. In fact, we think that questions of interactivity and non-interactivity are most likely to be settled in the traditional fashion, by experiments assessing whether manipulations of multiple variables do, in fact, result in interaction effects on the dependent measure of interest. To take just one example, an important issue in the reading literature is whether the two principal lexical effects on fixation durations in reading, frequency and predictability, have additive or interactive effects. On the basis of repeated experimental failures to demonstrate interactivity (e.g., Hand, Miellet, O'Donnell, & Sereno, 2010; Rayner, Ashby, Pollatsek, & Reichle, 2004; Slattery, Staub, & Rayner, 2011), both of the principal models of eye movements in reading (Engbert et al., 2005; Reichle et al., 2003) have explicitly adopted the assumption of additivity. But this is just one pair of variables (and one dependent measure, in one task), and it is likely that there are many theoretically important interactions to be found.

Despite the fact that we disagree with their substantive claims, we think that Stephen and Mirman (2010) make an important contribution. Though the eye movement literature has taken fixations and saccades to be central theoretical constructs, if it had turned out to be the case that distributions of gaze steps were well fit by any of the parameterizations tested by Stephen and Mirman, this might have provided an important alternative way of understanding eye movement behavior. It might have suggested that though the eyes appear to alternate between a stable state and a moving state, in fact this appearance is misleading, as the eyes actually engage in continuous movement, with the rate of movement varying as predicted by the best-fitting parameterization of the distance measure. The fact that the distribution of gaze steps is clearly a mixture, however, reaffirms the established fixations-and-saccades model of eye movement control. This is an important conclusion.

[4] To draw an analogy: While any useful description of automobile behavior will recognize that a car alternates between a stationary state and a moving state, it is surely the case that the time spent in the stationary state, and the distance traveled in the moving state, are each determined by complex interactions of variables.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrox & F. Caski (Eds.), *Second international symposium on information theory* (p. 267). Budapest: Akademiai Kiado.

Aks, D. J., & Sprott, J. C. (2003). Resolving perceptual ambiguity in the Necker Cube: A dynamical systems approach. *Journal of Nonlinear Dynamics & The Life Sciences, 7*, 159–178.

Aks, D. J., Zelinsky, G., & Sprott, J. C. (2002). Memory across eyemovements: 1/f Dynamic in visual search. *Journal of Nonlinear Dynamics & the Life Sciences, 6*, 1–25.

Benaglia, T., Chauveau, D., Hunter, D. R., & Young, D. (2009). mixtools: An R Package for Analyzing Finite Mixture Models. *Journal of Statistical Software, 32*(6), 1–29. <http://www.jstatsoft.org/v32/i06/>.

Boghen, C., Troost, B. T., Daroff, R. B., Dell'Osso, L. F., & Birkett, J. E. (1974). Velocity characteristics of normal human saccades. *Investigative Ophthamology and Visual Science, 13*, 619–623.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 84–107.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B, 39*(1), 1–38.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777–813.

Gilchrist, I. (2011). Saccades. In S. P. Liversedge, Iain. D. Gilchrist, & Stefan. Everling (Eds.), *The Oxford handbook of eye movements*. Oxford, UK: Oxford University Press.

Hand, C. J., Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2010). The frequency-predictability interaction in reading: It depends on where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1294–1313.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience, 2*, 194–203.

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: A neurophysiological analysis. *Trends in Neurosciences, 32*, 463–475.

Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology, 44*, 190–204.

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review, 117*, 382–405.

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org>.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446–461.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 720–732.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445–526.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*, 461–464.

Slattery, T. J., Staub, A., & Rayner, K. (2011). Saccade launch site as a predictor of fixation durations in reading: Comments on Hand, Miellet, O'Donnell, Sereno (2010). *Journal of Experimental Psychology: Human Perception & Performance*. doi:10.1037/a0025980.

Staub, A., Abbott, M., & Bogartz, R. S. (submitted for publication). Anticipatory eye movements in the visual world: Language influences where the eyes move, but not when.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations.

Journal of Experimental Psychology: Human Perception and Performance, 36, 1280–1293.

Stephen, D. G., & Mirman, D. (2010). Interactions dominate the dynamics of visual cognition. *Cognition, 115*, 154–165.

Stephen, D. G., Mirman, D., Magnuson, J. S., & Dixon, J. A. (2009). Lévy-like diffusion in eye movements during spoken-language comprehension. *Physical Review E, 79*, 056114.

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivey, J. C. (1995). Integration of visual and linguistic information in spoken-language comprehension. *Science, 268*, 1632–1634.

White, S. J., & Staub, A. (2011). The distribution of fixation durations during reading: Effects of stimulus quality. *Journal of Experimental Psychology: Human Perception & Performance*. doi: 10.1037/a0025338.

Yee, R. D., Schiller, V. L., Lim, V., Baloh, F. G., Baloh, R. W., & Honrubia, V. (1985). Velocities of vertical saccades with different eye movement recording methods. *Investigative Ophthalmology and Visual Science, 26*, 938–944.