Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Mathematical Psychology 48 (2004) 186–195
Journal of Mathematical Psychology
http://www.elsevier.com/locate/jmp

A note on the sampling properties of the Vincentizing (quantile averaging) procedure

Yi Jiang, Jeffrey N. Rouder,* and Paul L. Speckman

Department of Psychological Science, 210 McAlester Hall, University of Missouri, Columbia, MO 65211, USA

Received 3 February 2003; revised 14 January 2004

Abstract

To assess the effect of a manipulation on a response time distribution, psychologists often use Vincentizing or quantile averaging to construct group or "average" distributions. We provide a theorem characterizing the large sample properties of the averaged quantiles when the individual RT distributions all belong to the same location-scale family. We then apply the theorem to estimating parameters for the quantile-averaged distributions. From the theorem, it is shown that parameters of the group distribution can be estimated by generalized least squares. This method provides accurate estimates of standard errors of parameters and can therefore be used in formal inference. The method is benchmarked in a small simulation study against both a maximum likelihood method and an ordinary least-squares method. Generalized least squares essentially is the only method based on the averaged quantiles that is both unbiased and provides accurate estimates of parameter standard errors. It is also proved that for location-scale families, performing generalized least squares on quantile averages is formally equivalent to averaging parameter estimates from generalized least squares performed on individuals. A limitation on the method is that individual RT distributions must be members of the same location-scale family.

Keywords: RT analysis; Quantile regression; Vincentizing

1. Introduction

Response time (RT), the time taken to complete a task, is a common dependent variable in experimental psychology and has been used repeatedly to draw inferences about the nature of mental processing. Most researchers tend to analyze the mean RT, but a growing number of researchers are examining whole RT distributions as a means of providing more extensive and insightful tests of cognitive and perceptual theories (e.g., Ashby, Tien, & Balakrishnan, 1993; Dzhafarov, 1992; Hockley, 1984; Logan, 1992; Luce, 1986; Ratcliff, 1978; Ratcliff & Rouder, 1998, 2000, Rouder, 2000; Spieler, Balota, & Faust, 1996; Theeuwes, 1992, 1994; Townsend & Nozawa, 1995; Van Zandt, Colonius, & Proctor, 2000). For many paradigms and research questions, it is advantageous to use several participants and estimate a *group* response time distribution that reflects the distribution of an average or typical participant. To this end, it is imperative to have a set of statistically rigorous tools for forming group response time distributions.

Andrews and Heathcote's (2001) study on the effects of word frequency in identifying words is a recent example of constructing group RT distributions. Participants typically identify high-frequency words (e.g., cat) more quickly than low-frequency words (e.g., lynx), and this frequency advantage holds true for a variety of tasks such as naming, lexical decision, and deciding whether a word denotes an animate or inanimate object. Andrews and Heathcote's goal was to assess the loci of word frequency effects. They asked whether there was one common mechanism for word frequency effects that operated in all of these tasks. Their approach was to construct group RT distributions. From these group distributions, Andrews and Heathcote estimated parameters related to the "average" location and skew. Based on the patterns of changes in these parameters, Andrews and Heathcote were able to make detailed statements about loci of word frequency effects across a number of tasks.

*Corresponding author. Fax: +573-882-7710.
E-mail address: jeff@banta.psyc.missouri.edu (J.N. Rouder).

0022-2496/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmp.2004.01.002

The current state-of-the art in forming group response time distributions is a non-parametric technique called Vincentizing (Vincent, 1912). The goal of this paper is to provide a theorem about the sampling distributions of parameters obtained from Vincentizing. There are actually two different variants in the literature that are referred to as Vincentizing (Heathcote, 1996; Van Zandt, 2000). The first is the technique of averaging quantiles[1] across individual RT distributions to produce a group distribution. For example, the 10th percentile of the group distribution is the average of the 10th percentiles across all of the participants. This averaging procedure is applied to several quantiles until a reasonable group distribution is obtained. In the second variant, quantities called "Vincentiles" are averaged across participants. A Vincentile is the mean of observations between neighboring order statistics. For example, the 15% Vincentile is the mean of all observations between the 10th and 20th percentile. In practice, the two variants yield highly similar results, and occasionally authors do not specify which variant was used. We refer to the method of averaging either Vincentiles or quantiles across participants as Vincentizing. In this paper, we will be concerned exclusively with the variant of Vincentizing in which quantiles are averaged, where we can apply the well-known asymptotic distribution of sample quantiles. In contrast, we know of no statistical development about the distribution of Vincentiles.

Ratcliff (1979) provided a landmark study of Vincentizing. He found that the location, scale and shape of the group distribution roughly convey information about averaged location, scale, and shape of the individual distributions. Because of this property, Vincentizing has become a more frequently used tool in cognitive and perceptual psychology (e.g., Heathcote, Popiel, & Mewhort, 1991; Logan, 1992; Ratcliff & Rouder, 2000; Rouder, Ratcliff, & McKoon, 2000; Spieler, Balota, & Faust, 1996, 2000). The main advantage of Vincentizing is that it can be used in cases in which there are several individuals but only a few observations per individual. In contrast, it may be difficult if not impossible to estimate parameters from individuals with only a few observations. Vincentizing is therefore most useful in cases in which there are too few observations per individual to estimate each individual's distribution directly.

The main disadvantage of Vincentizing is the sampling properties of averaged quantiles are unknown. Without a theory of the sampling distribution of averaged quantiles, it is impossible to do inference on estimated group parameters without nonparametric or bootstrap methods. This deficit in the field motivates the current paper. In the following section we establish a theorem about averaged quantiles for *location-scale* distributions. The theorem implies that generalized least-squares regression (GLS) can be used to estimate the location and scale parameters of the group distribution. Generalized least squares also provides accurate estimates of parameter standard errors, which makes inference both possible and practical. Then, in the subsequent section, we show, through a small Monte Carlo study, that the quantile averaging/generalized least-squares method is both feasible and for large classes of models superior to other candidate methods including maximum likelihood. Finally, we show that for location-scale distributions, quantile averaging is unnecessary; the same parameter estimates can be obtained by fitting each individual's quantiles with GLS and averaging individual parameter estimates. This result is somewhat ironic. Vincentizing only yields consistent estimators for location-scale families (Thomas & Ross, 1980); yet, in this case it is unnecessary. We conclude with a discussion of when researchers can profitably use Vincentizing.

2. Statistical development

The goal in this section is to establish a theorem about the sample distributions of parameters estimated from averaged quantiles. We consider only location-scale families of random variables as it is only for these families that quantile averaging yields consistent parameter estimates (Thomas & Ross, 1980). First, we define a location-scale family of random variables; then we derive the sampling properties of averaged quantiles. With these sampling properties, it is relatively straightforward to provide an efficient means of estimating location and scale parameters as well as to provide an asymptotic result about their sampling distributions.

2.1. Location-scale distributions

Suppose each of several random variables, T_j, has a two-parameter distribution with parameters ψ_j and θ_j, $-\infty<\psi_j<\infty$, $\theta_j>0$. The collection is called a location-scale family if the distribution of $V_j=(T_j-\psi_j)/\theta_j$ is the same for all j. We refer to the distribution of V as the standard distribution. Denote the cumulative distribution function (cdf) of T_j as $F(t;\psi_j,\theta_j)$ and the cdf of the standard distribution as G. It follows that the collection of T_j is a location-scale family if and only if

$$F(t;\psi_j,\theta_j)=G\left(\frac{t-\psi_j}{\theta_j}\right). \quad (1)$$

[1] The term "quantile" is used to generalize the notion of quartiles, quantiles, deciles, percentiles, etc. In general, given $0<p<1$, the pth quantile is defined to be a value Q such that fraction p of the population or data set lies at or below Q and fraction $1-p$ lies above Q. Thus the median is the quantile with $p=0.5$, the third quartile corresponds to $p=0.75$, etc.

Location-scale families generally have probability density functions (pdf), in which case the density of T_j has the form

$$f(t;\psi_j,\theta_j)=\frac{1}{\theta_j}g\left(\frac{t-\psi_j}{\theta_j}\right), \tag{2}$$

where g is the pdf of the standard distribution V.

Many common distributions are location-scale families. For example, the collection of all normal distributions is a location scale family; the mean is the location parameter and the standard deviation is the scale parameter. The collection of exponential distributions is an example of a scale family (with location set to zero); the scale parameter is the standard deviation (or inverse of the rate parameter). Although there are important families that are not location-scale, sometimes one of the parameters in these families can be regarded as fixed and the result is a location-scale family. One example is the three-parameter Weibull distribution with cdf

$$F(t;\psi,\theta,\beta)=1-\exp\left\{\left(\frac{t-\psi}{\theta}\right)^{\beta}\right\},$$
$$t>\psi \text{ and } 0 \text{ elsewhere}, \tag{3}$$

with location parameter $-\infty<\psi<\infty$ and scale and shape parameters $\theta,\beta>0$. In many cases, the shape parameter can be taken as a single value across all individuals, resulting in a location-scale family. We have found that a Weibull with a shape parameter set to 2.0 well-approximates RT data (Rouder, Sun, Speckman, Lu, & Zhou, 2003). Another example of a location-scale family is the three parameter gamma when the number of stages is held constant (perhaps at 2). However, the Ex-Gaussian (Hohle, 1965), a popular three-parameter random variable for response time, is not a location-scale distribution.

We assume that the data for each individual come from a common location-scale family, but each person has his or her own unique location and scale parameters. Let F_j denote the cdf for the jth person's RT. Then

$$F_j(t)=F(t;\psi_j,\theta_j)=G\left(\frac{t-\psi_j}{\theta_j}\right).$$

2.2. Quantiles and sample quantiles

Vincentizing consists of averaging select quantile statistics across individuals. To make these notions precise, we review notions of population and sample quantiles. If a distribution has cdf G, the quantile function $Q(p)$, defined for $0<p<1$, is the inverse of the cdf, $Q(p)=G^{-1}(p)$. If G is not one-to-one, the definition is taken to be $Q(p)=\inf\{x: G(x)\geqslant p\}$. In the cases of continuous distributions considered here, G will be one-to-one for $G(x)>0$. When $F(t;\psi,\theta)$ is a member of a location-scale family, it follows from (1) that

$$F^{-1}(p;\psi_j,\theta_j)=\psi_j+\theta_jQ(p), \tag{4}$$

where Q is the quantile function of the standard distribution, G. Then the pth quantile of person j, say $Q_j(p)=F_j^{-1}(p)$, satisfies

$$Q_j(p)=\psi_j+\theta_jQ(p). \tag{5}$$

As in Thomas and Ross (1980), the average of these theoretical quantiles across subjects satisfies

$$\frac{1}{m}\sum_{j=1}^{m}Q_j(p)=\bar{\psi}+\bar{\theta}Q(p), \tag{6}$$

where $\bar{\psi}=\frac{1}{m}\sum_{j=1}^{m}\psi_j$ and $\bar{\theta}=\frac{1}{m}\sum_{j=1}^{m}\theta_j$. But this is exactly the quantile function of the distribution with cdf $F^*(t)=F(t;\bar{\psi},\bar{\theta})$, which can thus be taken to be a *group distribution* in the sense that it is a kind of average distribution for the m individuals in the group. Thomas and Ross also showed that location-scale families are the only families of distributions with this averaging property.

In practice, this averaging procedure is applied to sample quantiles rather than theoretical quantiles. There is no unique definition of a sample quantile, but the large sample theory employed here applies for a number of different definitions. To be specific, let $T_1,\ldots,T_n$ denote independent observations from an RT distribution with cdf F, and let $T_{(1)}<\cdots<T_{(n)}$ denote the *ordered* responses. Following Hyndman and Fan (1996), a number of common definitions of the pth sample quantile have the form

$$\hat{Q}(p)=(1-\gamma)T_{(j)}+\gamma T_{(j+1)}, \tag{7}$$

where

$$\frac{j-s}{n}\leqslant p<\frac{j-s+1}{n} \tag{8}$$

for some integer $-\infty<s<\infty$ and $\gamma\in[0,1]$. For example, the sample median is usually defined to be $0.5\{T_{(0.5n)}+T_{(0.5n+1)}\}$ for n even and $T_{(0.5(n+1))}$ for n odd. The first case corresponds to $s=0$ and $\gamma=0.5$, while the second has $s=0.5$ and $\gamma=0$. We use the sample quantiles as defined in (7) in the following statistical development.

2.3. Sampling properties of averaged quantiles

Although Thomas and Ross's result is informative, it pertains to quantile averaging at the population level and not at the sample level. In this section, we derive the sampling distribution of sample averaged quantiles, which we express as a theorem. The theorem leads, almost directly, to an efficient means of parameter estimation. Before stating the theorem and its consequences, we provide needed preliminary notation.

Preliminary notation. In practice, a researcher must first choose a set of quantiles (or percentiles). For example, Rouder et al. (2000) chose the 10th, 30th, 50th,

70th, and 90th percentiles, that is the quantiles defined for $p = 0.1, 0.3, 0.5, 0.7$, and 0.9. In general, we assume Vincentizing based on K quantiles with $0 < p_1 < \cdots < p_K < 1$. Let $\hat{Q}_j(p_k)$ denote the kth sample quantile for the jth individual. Let $Y(p_k)$ denote the corresponding averaged (i.e. Vincentized) quantile, that is,

$$Y(p_k) = \frac{1}{m}\sum_{j=1}^{m} \hat{Q}_j(p_k). \tag{9}$$

In stating the theorem, let $\boldsymbol{Y}_n = (Y(p_1), \ldots, Y(p_K))'$ denote the vector of Vincentized quantiles when there are n observations per individual. The asymptotic mean of $\boldsymbol{Y}_n$ will be shown to be $\boldsymbol{Z\eta}$, where $\boldsymbol{\eta} = (\bar{\psi}, \bar{\theta})'$ is an unknown parameter vector and

$$\boldsymbol{Z} = \begin{pmatrix} 1 & Q(p_1) \\ \vdots & \vdots \\ 1 & Q(p_K) \end{pmatrix} \tag{10}$$

is a fixed, known $K \times 2$ matrix. The asymptotic covariance matrix is a scale multiple of

$$\boldsymbol{\Sigma} = \begin{pmatrix} \frac{p_1(1-p_1)}{g(Q(p_1))^2} & \frac{p_1(1-p_2)}{g(Q(p_1))g(Q(p_2))} & \cdots & \frac{p_1(1-p_K)}{g(Q(p_1))g(Q(p_K))} \\ \frac{p_1(1-p_2)}{g(Q(p_1))g(Q(p_2))} & \frac{p_2(1-p_2)}{g(Q(p_2))^2} & \cdots & \frac{p_2(1-p_K)}{g(Q(p_2))g(Q(p_K))} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_1(1-p_K)}{g(Q(p_1))g(Q(p_K))} & \frac{p_2(1-p_K)}{g(Q(p_2))g(Q(p_K))} & \cdots & \frac{p_K(1-p_K)}{g(Q(p_K))^2} \end{pmatrix}. \tag{11}$$

With this notation, it is now possible to express the theorem about the sampling distribution of averaged quantiles. In the statement of the theorem, we need a technical condition on the number of order statistics used in the definitions of the quantiles in (7) and (8). We assume that the integer s in the definition is bounded for all p_k and n. In many definitions, $|s| \leqslant 1$.

Theorem. *Assume* $g(Q(p_k)) > 0$ *and* $\hat{Q}(p_k)$ *satisfies* (7) *and* (8) *with* $|s| \leqslant r$ *for some fixed* $r < \infty$, $k = 1, \ldots, K$, $j = 1, \ldots, m$, *and all* n. *Then for fixed* m,

$$\sqrt{n}\{\boldsymbol{Y}_n - \boldsymbol{Z\eta}\} \xrightarrow{\mathscr{L}} N(\boldsymbol{0}, \sigma^2\boldsymbol{\Sigma}) \quad \text{as } n \to \infty, \tag{12}$$

where $\boldsymbol{0} = (0,0)^T$ *and*

$$\sigma^2 = \frac{1}{m^2}\sum_{j=1}^{m} \theta_j^2.$$

Before proving the theorem we discuss its implication, namely that we can use (12) as the approximate likelihood of the parameters $\bar{\psi}$ and $\bar{\theta}$. Inference for these parameters is then a straightforward application from multivariate normality, where by the Gauss Markov theorem, generalized least squares provides efficient estimates (e.g., Christensen, 1996) in the following way.

We treat $\boldsymbol{Y}$ as if it has a multivariate normal distribution with mean $\boldsymbol{Z\eta}$ and variance–covariance matrix $\tau\boldsymbol{\Sigma}$, where $\tau = \sigma^2/n$ is an unknown parameter. Under these assumptions, the usual estimate (best linear unbiased estimate) of $\boldsymbol{\eta}$ is

$$\hat{\boldsymbol{\eta}} = (\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, \tag{13}$$

and the distribution of $\hat{\boldsymbol{\eta}} = (\hat{\psi}, \hat{\theta})'$ is approximately $N(\boldsymbol{\eta}, \tau(\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z})^{-1})$. Moreover,

$$\begin{aligned} \hat{\tau} &= (\boldsymbol{Y} - \boldsymbol{Z}\hat{\boldsymbol{\eta}})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \boldsymbol{Z}\hat{\boldsymbol{\eta}})/(K-2) \\ &= \boldsymbol{Y}'\{\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\}\boldsymbol{Y}/(K-2) \end{aligned} \tag{14}$$

is an unbiased estimate of τ. Letting

$$\boldsymbol{S} = [s_{ij}] = \hat{\tau}(\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z})^{-1}, \tag{15}$$

approximate confidence intervals for $\bar{\psi}$ and $\bar{\theta}$ are given by $\hat{\psi} \pm t_{\alpha/2,K-2}\sqrt{s_{11}}$ and $\hat{\theta} \pm t_{\alpha/2,K-2}\sqrt{s_{22}}$, respectively.

This result is closely related to the method introduced by Mosteller (1946) for "inefficient" estimates of parameters of the normal distribution based on selected order statistics from a single sample. Mosteller's work has been extended by a number of authors to various distributions. These authors also consider optimal choice of quantiles; see Eubank (1981) for further discussion and references.

Proof of Theorem 1. We begin with quantiles defined by (7) for $s = 0$ and $\gamma = 1$, that is, $\hat{Q}_j(p)$ is the $\lceil pn \rceil$ order statistic. By the asymptotic limit theorem for sample quantiles (see for example Sen & Singer, 1993, p. 169, Theorem 4.4.3), the jth person's quantiles have joint asymptotic distribution

$$\sqrt{n}\begin{pmatrix} \hat{Q}_j(p_1) - Q_j(p_1) \\ \vdots \\ \hat{Q}_j(p_K) - Q_j(p_K) \end{pmatrix} \xrightarrow{\mathscr{L}} N(\boldsymbol{0}, \tilde{\boldsymbol{\Sigma}}_j) \quad \text{as } n \to \infty, \tag{16}$$

where $\tilde{\boldsymbol{\Sigma}}_j$ is the $K \times K$ matrix with (r,s) element

$$\frac{p_r(1-p_s)}{f_j\{Q_j(p_r)\}f_j\{Q_j(1-p_s)\}}.$$

Using (2) and (5), this covariance term simplifies to

$$\frac{\theta_j^2 p_r(1-p_s)}{g\{Q(p_r)\}g\{Q(1-p_s)\}},$$

giving the asymptotic covariance $\theta_j^2\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is defined in (11). Consequently, Slutsky's Theorem allows us to average over j to obtain

$$\begin{aligned} &\sqrt{n}\begin{pmatrix} m^{-1}\{\sum_{j=1}^m \hat{Q}_j(p_1) - Q_j(p_1)\} \\ \vdots \\ m^{-1}\{\sum_{j=1}^m \hat{Q}_j(p_K) - Q_j(p_K)\} \end{pmatrix} \\ &\quad \xrightarrow{\mathscr{L}} N\left(\boldsymbol{0}, \frac{\sum_j \theta_j^2}{m^2}\boldsymbol{\Sigma}\right) \quad \text{as } n \to \infty. \end{aligned} \tag{17}$$

The theorem for this case now follows from (6) and (9).

To extend this result to arbitrary s and γ as defined in (7) and (8), we need a simple extension of (16). Because we are not aware of a reference, a short proof is included here. Again let $T_{(1)}, \ldots, T_{(n)}$ denote the order statistics for person j. It is well known (e.g. Ferguson, 1996) that these order statistics have the same joint distribution as $Q_j(U_{(1)}), \ldots, Q_j(U_{(n)})$, where $U_{(1)}, \ldots, U_{(n)}$ are the order statistics from a random sample of uniformly distributed random variables on [0, 1]. Furthermore, $U_{(i)} \stackrel{\mathcal{D}}{=} X_i / \sum_{j=1}^{n} X_i$, $i = 1, \ldots, n$, where the X_i are independent exponentially distributed (Gamma(1, 1)) random variables. For fixed integer $r > 0$, it follows that

$$n(U_{(\lceil pn \rceil + r)} - U_{(\lceil pn \rceil)}) \stackrel{\mathcal{D}}{=} \frac{n \sum_{i=1}^{r} X_i}{\sum_{j=1}^{n} X_j} \stackrel{\mathcal{D}}{\to} \text{Gamma}(r, 1).$$

Consequently, $\sqrt{n}(U_{(\lceil pn \rceil + r)} - U_{(\lceil pn \rceil)}) \to 0$ in probability. Using the delta method, the assumption on g, and the fact that $\hat{Q}_j(p) \to Q_j(p)$ in probability, it follows that

$$\sqrt{n}(T_{(\lceil pn \rceil + r)} - T_{(\lceil pn \rceil)}) \to 0$$

in probability as well. A similar results holds for fixed $r < 0$. Now choose $r < \infty$ so that $|s| < r$ for all n in (8) (s can depend on n). Then using (7) again,

$$\sqrt{n}|\hat{Q}_j(p) - T_{(\lceil pn \rceil)}| \leqslant \sqrt{n}\{(1-\gamma)|T_{(j)} - T_{(\lceil pn \rceil)}| + \gamma|T_{(j+1)} - T_{(\lceil pn \rceil)}|\}.$$

There are several possibilities. If $j \leqslant \lceil np \rceil \leqslant j + 1$, the right side is bounded by

$$\sqrt{n}(|T_{(\lceil pn \rceil - r)} - T_{(\lceil pn \rceil)}| + |T_{(\lceil pn \rceil + r)} - T_{(\lceil pn \rceil)}|) \to 0$$

in probability. The same bound holds for $s + 1 < \lceil pn \rceil$ or $s > \lceil pn \rceil$. Thus convergence holds uniformly over the finite set of quantiles $p_1, \ldots, p_K$, and the general case of the theorem follows from another application of Slutsky's theorem.

3. Application to the Weibull distribution

To illustrate the implications of the above theorem, we demonstrate its application to a Weibull distribution (3) with fixed shape parameter β. The three-parameter Weibull is a flexible form that has been used by several researchers (e.g., Chechile, 2001; Logan 1992; Hale, Myerson, Robertson & Zar-Kessler, 2002; Rouder et al., 2003). The three-parameter Weibull is not a location-scale family. To make such a family, we fixed the shape to a constant β. We have found that for college-aged students performing simple tasks, the shape parameter of the Weibull is around $\beta = 2$. This value also agrees well with the reported fits of Logan (1992) and is used by Chechile (2001). With this fixed-shape Weibull, we can apply the theorem. The degree to which the fixed shape is valid or nearly valid is a function of the task. It may be correct or near correct for homogeneous groups doing relatively simple tasks. It may break down in tasks in which there may be differing strategies across individuals (e.g., solving crossword puzzles).

The Weibull distribution (3) with fixed shape $\beta > 0$ has base cdf $G(v) = 1 - e^{-v^\beta}$, $v \geqslant 0$. Thus $g(v) = \beta v^{\beta-1} e^{-v^\beta}$, $v > 0$, and

$$Q(p) = \{-\log(1-p)\}^{1/\beta}. \tag{18}$$

As a result,

$$g(Q(p)) = \beta Q(p)^{\beta-1} \exp\{-Q(p)^\beta\} = \beta(1-p)\{-\log(1-p)\}^{(\beta-1)/\beta}. \tag{19}$$

For the special case $\beta = 2$,

$$g\{Q(p)\} = 2(1-p)\{-\log(1-p)\}^{1/2}.$$

With these expressions, (13) can be used to estimate the group location and scale parameters. Computations for generalized least squares estimation can be carried out using the Splus statistics package (Insightful Corp.) or the public domain R statistics program, and code is available at `wwww.missouri.edu/~pcl`.

If β is unknown, it can be estimated using the approximate likelihood of the vector of Vincentized quantiles $\boldsymbol{Y}$. Let $\boldsymbol{Z}(\beta)$ be given by (10) with Q in (18), and let $\boldsymbol{\Sigma}(\beta)$ be given by (11) using (19). Setting $\tau = \sum_{j=1}^{m} \theta_j^2 / (nm^2)$, the approximate log likelihood of $\boldsymbol{Y}$ is equivalent to

$$-\frac{\tau}{2}(\boldsymbol{Y} - \boldsymbol{Z}(\beta)\boldsymbol{\eta})' \boldsymbol{\Sigma}(\beta)^{-1} (\boldsymbol{Y} - \boldsymbol{Z}(\beta)\boldsymbol{\eta}) - \frac{K}{2}\log\tau - \frac{1}{2}\log|\boldsymbol{\Sigma}(\beta)|.$$

After maximizing in $\boldsymbol{\eta}$ and τ using (13) and (14), the likelihood is equivalent to

$$-\frac{K}{2}\log(\boldsymbol{Y}'[\boldsymbol{\Sigma}(\beta)^{-1} - \boldsymbol{\Sigma}(\beta)^{-1}\boldsymbol{Z}(\beta)\{\boldsymbol{Z}(\beta)'\boldsymbol{\Sigma}(\beta)^{-1}\boldsymbol{Z}(\beta)\}^{-1} \boldsymbol{Z}(\beta)'\boldsymbol{\Sigma}(\beta)^{-1}]\boldsymbol{Y}) - \frac{1}{2}\log|\boldsymbol{\Sigma}(\beta)|.$$

Maximizing this expression over $\beta > 0$ gives an estimate of the common scale parameter to use for the data set.

4. Simulation study

To test the behavior of the quantile averaging/generalized least-squares method with finite samples, we performed a small Monte Carlo study in which the estimates from a laboratory data set served as "true values" for location and scale parameters.[2] Group true

[2] Data come from the calibration phase of an implicit learning experiment performed by Michael Stadler. Participants were given one of four simple stimuli (an asterisk at one of four locations on a computer screen) and had to make one of four responses by pressing a key on a computer keyboard. Each stimulus was mapped to a unique response. Eighty participants responded to 120 such trials. Rouder et al. (2003) used MLE to estimate Weibull location, scale, and shape parameters. In this paper, we use Rouder et al.'s MLE location and

values were obtained by averaging all the individuals' true values. Simulated data consisted of 10,000 replicates of 20 observations for each of 80 participants. Data were analyzed separately for each replicate.

There is one remaining issue to implement the generalized least-squares method: the algorithm for finding sample quantiles. Eqs. (7) and (8) are broad and include many standard algorithms. Surprisingly, the form of this algorithm will have a noticeable effect on parameter estimation. Hence, the issue must be addressed with care and we document two methods. A fuller treatment of quantile estimators is given by Hyndman and Fan (1996).

In the simulation study, we took $\mathbf{p} = (0.1, 0.3, 0.5, 0.7, 0.9)$. This is a fairly conventional set that has been previously used (e.g., Ratcliff & Rouder, 2000; Rouder et al., 2000). Given that there are 20 observations per person, one approach is to average the 2nd, 6th, 10th, 14th and 18th ranked observation. This uses the naive rule of taking ranks $20p_i$ for $i = 1, \ldots, 5$. But, these quantile estimators are systematically biased low. We used the following procedure to get approximately unbiased quantile estimators. Let $V_{(1)} < \cdots < V_{(n)}$ denote the order statistics of a sample of size n from the base distribution G, and let $q_{i:n} = E(V_{(i)})$ be the expected value of the ith order statistic. (This is the "score" used in nonparametric statistics.) For the Weibull distribution, we estimated these scores by a simple Monte Carlo procedure. We took 100,000 random samples of size $n = 20$ from the base Weibull distribution with $\psi = 0$ and $\theta = 1$, and we ordered each sample. We then calculated the average of each order statistic. The result was $q_{2:20} \approx 0.3007$, etc. We then computed $p_i = G(q_{i:n})$, $i = 1, \ldots, K$, to obtain the quantiles estimated for which the chosen order statistics were unbiased. The result for the ranks $(2, 6, 10, 14, 18)$ was $p = (0.0867, 0.2825, 0.4789, 0.6756, 0.8722)$. To get more accurate estimates of the group parameters $(\bar{\psi}, \bar{\theta})$, one can replace the nominal values of $\mathbf{p}$ by these new estimates.

This calculation of exact target probabilities $\boldsymbol{p}$ is appropriate when all persons in the study have the same number of responses. In cases where this is not true, we explored two alternative methods for obtaining less biased quantile estimators. The first method is commonly used in statistics packages (e.g. PROC UNIVARIATE in SAS). Suppose $p(n+1) = i + c$, where i is an integer and $0 \leqslant c < 1$. Then take $\hat{Q}_j(p) = (1-c)T_{(i),j} + cT_{(i+1),j}$, where $T_{(1),j} < \cdots < T_{(n),j}$ are the order statistics for person j. The estimator is unbiased when sampling from a uniform distribution. For example, with $n = 20$, the quantile corresponding to $p = 0.1$ has rank $0.1(21) = 2.1$. The resulting sample quantile for $p = 0.1$ is $0.9\,Y_2 + 0.1\,Y_3$. Quantiles from this method are termed "uniform quantiles" in the upcoming analysis since they are unbiased estimators for the uniform distribution. The second method again uses simulation to get virtually unbiased quantile estimators. As above, use Monte Carlo simulation to estimate the scores $q_{1:n}, \ldots, q_{n:n}$ for the base distribution G. If i is the integer such that $q_{i:n} \leqslant Q(p) < q_{(i+1):n}$, let $c = (Q(p) - q_{i:n})/(q_{(i+1):n} - q_{i:n})$ and take $\tilde{Q}_j(p) = (1-c)T_{(i),j} + cT_{(i+1),j}$. For example, suppose $n = 20$ and $p = 0.1$. Using the fixed shape Weibull in (18) with $\beta = 2$, $Q(0.1) = 0.3246$. From simulation based on 100,000 trials, $q_{2:20} \approx 0.3007$, and $q_{3:20} \approx 0.3809$, so $c \approx 0.298$, and the estimated quantile is $\tilde{Q}_j(0.1) = 0.702T_{(2),j} + 0.298T_{(3),j}$. Quantiles from this method will be termed "unbiased quantiles". This method can be used for each individual separately if there are unequal numbers of observations. It can be shown that the asymptotics leading to the theorem also hold for both of these modified definitions of sample quantiles.

We also tested two different estimation alternatives: ordinary least squares based on the approximate model $E(\boldsymbol{Y}_n) = \boldsymbol{Z\eta}$ from the theorem and an ML-based method. Since OLS is simpler to implement than generalized least squares, we were interested in comparing the results with those obtained by the optimal method. We were also interested in comparing the results based on quantile averaging with direct application of maximum likelihood to each person's data. In the ML-based method, each individual's location and scale parameters were directly estimated with MLE.[3] Then, these individual estimates were averaged to produce group estimates. ML has theoretical optimal properties, at least for large samples from "regular" distributions. Thus, given all the data, it might be surmised that the ML-based method would be better than any method following reduction to a few estimated quantiles. However, the usual large sample optimality results do not apply to Weibull distributions. The problem is that the location parameter defines the lower bound on the distribution.

For each replicate, we calculated estimation errors—the signed difference between an estimate and the "true" value. Fig. 1 shows histograms of the estimation errors for four different methods. The top row shows results from generalized least squares with the unbiased quantiles. These appear to have little bias and are the most efficient (i.e. smallest variance) estimates among those tested. In fact, the mean error in location and scale is less than 0.07 ms (this small amount is not statistically different from 0). Perhaps surprisingly, the OLS method

(*footnote continued*)
scale parameter estimates as "true values". Further details about the data-set and the hierarchical model can be found in Rouder et al. We are grateful to Michael Stadler for use of his data.

[3] The negative log-likelihood function was minimized with repeated applications of the simplex algorithm (Nelder & Mead, 1965).

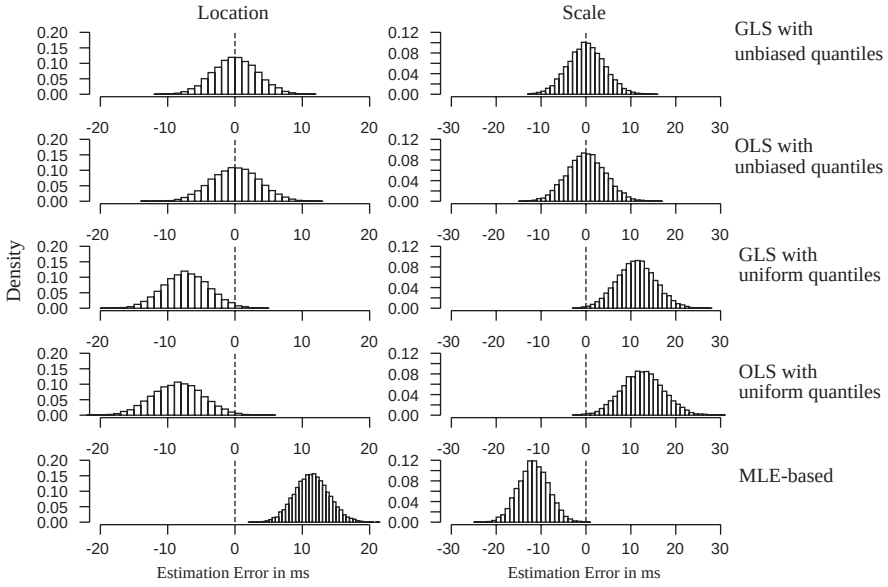

Fig. 1. Histograms of the estimation error. The top two rows show errors for the averaged-quantile with generalized least-squares estimation. The third row shows errors with ordinary least-squares estimation while the bottom panel row errors with the ML-based method.

Table 1
Variability in estimation*

Method	Location Estimated std. dev.	Location average std. error	Location Coverage prob.	Scale Estimated std. dev.	Scale average std. error	Scale Coverage prob.
Five quantiles						
GLS	3.29	2.85	0.942	4.08	3.56	0.940
OLS	3.57	1.97	0.849	4.29	2.04	0.796
Eighteen quantiles						
GLS	3.27	3.17	0.948	4.04	3.95	0.948
OLS	3.75	1.14	0.466	4.36	1.22	0.439

*Reported in milliseconds

with unbiased quantiles did virtually as well as the GLS method; it had no detectable bias and was only slightly more variable. When the uniform quantiles were used, there was noticeable bias on the order of 10 ms. Once again, the histogram of OLS estimates was virtually identical to that of GLS estimates. As can be seen, the algorithm for estimating quantiles affected the bias in the parameter estimates.

The most biased method was the ML-based method, which resulted in bias of about 12 ms. Although this was the largest bias of all methods, we were actually surprised by this small degree of bias and relative efficiency. The reason for our surprise is that ML-based estimation is virtually impossible with the three-parameter Weibull with 20 observations per participant (Smith & Naylor, 1987; Rouder et al., 2003; Rouder & Speckman, in press). The degree of instability with the three-parameter Weibull typically leads to a massive instability—parameter estimates vary by several orders of magnitude. Our simulation shows that estimation with the fixed-shape Weibull is far more feasible and not subject to this type of massive instability. Hence, the bias present in the ML estimates is not caused by instability in numerical estimation and reflects the simple fact that ML is somewhat biased at these small sample sizes.

The histograms show the accuracy of estimation. Researchers concerned with inference need to be able to estimate standard errors and construct confidence intervals as well. The standard errors for the parameter estimates are derived from (14). We checked on the accuracy of this approximation in the simulation. Eq. (14) was used to calculate standard error estimates on each replicate of the simulation. The average over all simulations is shown in the column labeled "Average std. error" in Table 1. The goal is to see whether this estimate provides a reasonable approximation to the underlying standard deviation of the parameters. The sampling standard deviation of each parameter estimate can be estimated from the simulation (it is the standard deviation of the histograms), and the results are shown in the column labeled "Estimated std. dev." The results

for the current simulation for five quantiles are shown under the label of "Five quantiles". Only the GLS and OLS methods with unbiased quantiles are included in the table as these are the only methods that yielded parameter estimates with negligible bias.

For GLS, estimates from (14) modestly underestimate the true values. But, for OLS, the underestimation is more severe. Underestimation of variability is quite problematic as it tends to an underestimation of confidence intervals and a greater likelihood of Type I errors.

To assess how underestimation of standard errors affects inference, we computed 95% t-interval confidence intervals for $\bar{\psi}$ and $\bar{\theta}$ for each replicate, and we recorded the proportion of times the confidence intervals contained the true parameters. The result is labeled "Coverage prob." in Table 1. Coverage is critical for inference and if the coverage is too low, then the researcher has an increased chance of mis-localizing the true parameter. For GLS, the coverages are between 0.94 and 0.942 for nominal 95% confidence levels with three degrees of freedom. (This means that the standard errors calculated with the normal approximation are quite accurate, even with samples as small as 20 per subject.) The same is not true for OLS. The more severe underestimation of parameter variability leads to a much lower coverage rate, between 0.796 and 0.840 for nominal 95% confidence intervals. This low coverage translates to an increased actual Type I error rate over the nominal 0.05 significance level—as high as 0.20 in this application. Researchers cannot trust inference based on OLS.

The reason that OLS underestimates parameter variability is that it fails to account for correlation in the quantiles. It was initially hoped that the probability targets spacing was sufficiently great to reduce substantial correlation, but this is not the case. To more forcefully demonstrate this divergence between GLS and OLS, we performed a final simulation to maximize the correlation between quantiles by choosing as many target probabilities as possible. We quantile averaged the 2nd through 19th ranked observations. Using results in Csörgő (1983), the Theorem can be extended to cover this case as well. We then simulated the associated probability targets and performed GLS and OLS estimation. Once again, both estimators are virtually unbiased. In fact, mean errors were on the order of a few microseconds. Table 1 shows standard error estimates from both (14) and the simulations directly. The same pattern is evident with slight underestimation of variance for GLS and severe underestimation for OLS. The coverage proportions were also computed. For GLS, the coverage was around 0.948 for both parameters. For OLS, the increased correlation resulted in coverage rates less than 0.50. That is, for OLS there is an actual Type I error over 0.50 for a nominal significance level of 0.05. In our opinion, the results for GLS are very encouraging. They suggest that the asymptotics work well even with as few as 20 observations per individual.

5. Quantile averaging vs. parameter averaging

The theorem shows that generalized least squares is an efficient means of estimating location and scale parameters from sample quantiles. In the previous application, we applied the theorem to averaged quantiles, but it can be applied to estimate quantiles from individual participants as well. The parameters estimated from individual's sample quantiles can then be averaged to yield estimates of $(\bar{\psi}, \bar{\theta})$. Which method gives better group parameter estimates, averaging each person's estimated parameters or estimation based on averaged quantiles? The Theorem shows that GLS is (at least asymptotically) the most efficient method based on the average quantiles. Perhaps surprisingly, the GLS method can also be viewed as a parameter averaging method. To see this, let $\boldsymbol{X}_j = (\hat{Q}_j(p_1), \ldots, \hat{Q}_j(p_K))$ denote the estimated quantiles for person $j = 1, \ldots, m$. Then

$$\boldsymbol{Y} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{X}_i$$

and consequently from (13)

$$\hat{\boldsymbol{\eta}} = \frac{1}{m}\sum_{i=1}^{m} \hat{\boldsymbol{\eta}}_i \quad \text{where } \hat{\boldsymbol{\eta}}_i = (\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_i.$$

In other words, $\hat{\eta}$ is the average of the estimates $\hat{\boldsymbol{\eta}}_i = (\hat{\psi}_i, \hat{\theta}_i)'$, which are calculated by GLS for each individual in the sample. (The individual estimates are the quantile-based estimates studied, for example, by Eubank (1981).) That is, for the GLS method, averaging parameters estimated from individual's quantiles or estimating parameters from averaged quantiles yield identical estimates for location-scale families. This result strikes us as ironic. It is only for location-scale that quantile averaging yields consistent estimators. Yet, in this case, quantile averaging offers no advantage over averaging individual's GLS estimates. Thus, in the context of location-scale families, the issue is not whether quantile averaging is better than parameter averaging, it is simply how best to estimate individual parameters.

Maximum likelihood is widely known to have good properties (best possible in an asymptotic sense) under suitable regularity conditions. In fact, the MLE of $(\bar{\psi}, \bar{\eta})$ is the average of the individual MLEs. For certain distributions, it is known to be optimal. For example, if the RT distributions are exponential with $\psi_j = c$ fixed before hand, MLE is known to be minimum variance unbiased (see e.g. Lehmann, 1991), hence parameter averaging will be also. In this case, averaging MLEs will

be more efficient than averaging the GLS quantile estimates of the Theorem. However, our simulation shows that averaging MLEs is not necessarily the best procedure for small samples with the Weibull distribution. In fact, because the support of the three-parameter Weibull with shape fixed depends on the location parameter, the regularity conditions for the oft-cited optimality of MLE do not hold. Moreover, the asymptotic justification for the MLE does not always result in good small sample properties. Thus our simulation shows that it is worthwhile to investigate alternative estimators, at least for modest sized samples.

6. General discussion

We have presented an asymptotic result allowing researchers to calculate estimated standard errors for Vincentized estimates using averaged quantiles. These results also suggest that generalized least squares is an efficient means of estimation and inference for large samples per individual. This efficiency is supported with Monte Carlo simulations. Perhaps surprisingly, the results appear to hold for samples even as small as 20 per individual. Competing methods, either ML-based or OLS, were inferior in that they were either biased or had poor properties for inference.

Although quantile averaging with GLS is a suitable means for estimation of and inference about grouped parameters in location-scale models of response time, it is not necessary. For the general least squares approach presented here, parameter averaging individual estimates yields the same exact estimates. This result is ironic. Location-scale families are the only ones in which quantile averaging yields consistent estimators, yet, even here, quantile averaging is not superior to averaging individual parameter estimates.

Although Vincentizing is unnecessary with location-scale families, it may still be useful in practice with non-location-scale family distributions, especially if there are a few samples per individual and the distribution is irregular. Rouder and Speckman (in press), for example, show through Monte Carlo simulation that Vincentizing is superior to averaging individual parameters for estimating group-average parameters from a three-parameter Weibull distribution. Yet, it is inferior to parameter averaging for the ex-Gaussian or shifted Wald distributions. Researchers who use Vincentizing with non-location-scale families should bear in mind that the estimators are inconsistent. Although the efficiency of a Vincentized estimate may be superior to a consistent method for small samples, asymptotically, the consistent method with be superior. Therefore, we recommend that Vincentizing be used sparingly and only in conjunction with simulations that show its clear advantage for a given distribution at a given sample size.

Vincentizing is a classical method. There are newer computational methods that allow researchers the practical ability to model individuals and treatments simultaneously. Particularly promising examples include Bayesian Hiearchical Modeling (e.g., Gelman, Carlin, Stern, & Rubin, 1995; RT applications by Rouder et al., 2003) and proportional hazard function modeling (Therneau & Grambsch, 2000; RT applications by Wenger, Schuster, Petersen, & Petersen, in press). However, in situations where Vincentizing is adequate, the ability to compute approximate standard errors will allow researchers to perform inference and hence better interpret their data.

Acknowledgments

We are grateful to Michael Stadler who allowed us use of his data and to Andrew Heathcote who provided helpful comments. This research is supported by National Science Foundation Grant SES-0095919 to J. Rouder, D. Sun, and P. Speckman, University of Missouri Research Board Grant 00-77 to J. Rouder, and National Science Foundation grant DMS-9972598 to D. Sun and P. Speckman.

References

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 514–544.

Ashby, F. G., Tien, J.-Y., & Balikrishnan, J. D. (1993). Response time distributions in memory scanning. *Journal of Mathematical Psychology*, *37*, 526–555.

Chechile, R. A. (2001). New procedures for individual RT analysis. *Presented at 42nd annual meeting of the psychonomic society*, Orlando, FL.

Christensen, R. (1996). *Plane answers to complex questions: The theory of linear models* (2nd ed.). Berlin: Springer.

Csörgő, M. (1983). *Quantile processes with statistical applications*. Philadelphia: Society for Industrial and Applied Mathematics.

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology*, *36*, 235–268.

Eubank, R. L. (1981). A density-quantile function approach to optimal spacing selection. *The Annals of Statistics*, *9*, 494–500.

Ferguson, T. S. (1996). *A course in large sample theory*. London: Chapman & Hall.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. London: Chapman & Hall.

Hale, S., Myerson, J. A., Robertson, S., & Zar-Kessler, C. (2002). Response time distributions and conditional accuracy functions of younger and older adults performing a same-different task. *Presented at cognitive aging conference*, Atlanta, GA.

Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Behavioral Research Methods, Instruments, and Computers*, *28*, 427–445.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, *109*, 340–347.

Hockley, W. E. (1984). Analysis of reaction time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 598–615.

Hohle, R. H. (1965). Inferred components of reaction time as a function of foreperiod duration. *Journal of Experimental Psychology*, *69*, 382–386.

Hyndman, R. J., & Fan, Y. N. (1996). Sample quantiles in statistical packages. *American Statistician*, *50*, 361–365.

Lehmann, E. L. (1991). *Theory of point estimation*. Belmont, CA: Wadsworth.

Logan, G. D. (1992). Shapes of reaction time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 883–914.

Luce, R. D. (1986). *Response times*. New York: Oxford University Press.

Mosteller, F. (1946). On some useful ''inefficient'' statistics. *Annals of Mathematical Statistics*, *17*, 377–408.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308–313.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

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

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for decisions between two choices. *Psychological Science*, *9*, 347–356.

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model analysis of letter masking. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 127–140.

Rouder, J. N. (2000). Assessing the roles of change discrimination and luminance integration: Evidence for a hybrid race model of perceptual decision making in luminance discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 359–378.

Rouder, J. N., Ratcliff, R., & McKoon, G. M. (2000). A neural network model of priming in object recognition. *Psychological Science*, *11*, 13–19.

Rouder, J. N., & Speckman, P.L. (in press). A note on the advantages and disadvantages of the Vincentizing procedure. *Psychonomic Bulletin and Review*.

Rouder, J. N., Sun, D., Speckamn, P. L., Lu, J., & Zhou, D. (2003). A hierarchical Bayesian statistical framework for skewed variables with an application to response time distributions. *Psychometrika*, *68*, 587–604.

Sen, P. K., & Singer, J. M. (1993). *Large sample methods in statistics. An introduction with applications*. London: Chapman & Hall.

Smith, R. L., & Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistician*, *36*, 358–369.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 461–479.

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 506–526.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 799–806.

Therneau, T., & Grambsch, P. M. (2000). *Modeling survival data: Extending the Cox model*. New York: Springer.

Thomas, E. A. C., & Ross, B. (1980). On appropriate procedures for combining probability distributions within the same family. *Journal of Mathematical Psychology*, *21*, 136–152.

Townsend, J. T., & Nozawa, G. (1995). On the spatio-temporal properties of elementary perception: An investigation on parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*, 321–359.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin and Review*, *7*, 424–465.

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin and Review*, *7*, 208–256.

Vincent, S. B. (1912). The function of vibrissae in the behavior of the white rat. *Behavioral Monographs*, *1* (Whole No. 5).

Wenger, M. T., Schuster, C., Petersen, L. M., & Petersen, R. C. (in press). Application of proportional hazard models to reaction time data. In C. Bergeman, & S. Boker (Eds.), *Notre-Dame series on quantitative methodologies*. Mahwah, NJ: Erlbaum.