qPR: An adaptive partial-report procedure based on Bayesian inference

Jongsoo Baek Yonsei Institute of Convergence Technology, Yonsei University, Incheon, Korea

Luis Andres Lesmes Adaptive Sensory Technology, Inc, Boston, MA, USA

Zhong-Lin Lu Laboratory of Brain Processes (LOBES), Center for Cognitive and Brain Sciences & Center for Cognitive and Behavioral Brain Imaging, Department of Psychology, The Ohio State University, Columbus, OH, USA

Iconic memory is best assessed with the partial report procedure in which an array of letters appears briefly on the screen and a poststimulus cue directs the observer to report the identity of the cued letter(s). Typically, 6–8 cue delays or 600–800 trials are tested to measure the iconic memory decay function. Here we develop a quick partial report, or qPR, procedure based on a Bayesian adaptive framework to estimate the iconic memory decay function with much reduced testing time. The iconic memory decay function is characterized by an exponential function and a joint probability distribution of its three parameters. Starting with a prior of the parameters, the method selects the stimulus to maximize the expected information gain in the next test trial. It then updates the posterior probability distribution of the parameters based on the observer's response using Bayesian inference. The procedure is reiterated until either the total number of trials or the precision of the parameter estimates reaches a certain criterion. Simulation studies showed that only 100 trials were necessary to reach an average absolute bias of 0.026 and a precision of 0.070 (both in terms of probability correct). A psychophysical validation experiment showed that estimates of the iconic memory decay function obtained with 100 qPR trials exhibited good precision (the half width of the 68.2% credible interval = 0.055) and excellent agreement with those obtained with 1,600 trials of the conventional method of constant stimuli procedure (RMSE = 0.063). Quick partial-report relieves the data collection burden in characterizing iconic memory and makes it possible to assess iconic memory in clinical populations.

Introduction

It has been known that human memory is composed of three substorages: sensory memory, short-term memory, and long-term memory (Atkinson & Shiffrin, 1968; Baddeley & Hitch, 1974; James, 1890; Sperling, 1963). Sensory memory is the literal, modality-specific neural representation of sensory stimuli in the human brain (Sperling, 1960). Sensory inputs from the environment are initially stored in sensory memory and processed in subsequent stages of perception and cognition. The existence of sensory memory was first demonstrated in Sperling's seminal partial-report experiment in which three rows of letters were briefly displayed, and observers were asked to report the identity of the letters in one of the rows cued with a high, middle, or low frequency tone after a variable delay (Figure 1a). The number of correctly reported items in the partial-report condition was compared with that in the whole-report condition, in which observers were asked to report all items in the display. Figure 1b illustrates results from a typical partial-report experiment. The number of correctly reported letters decreases rapidly with increasing cue delay in partial-report, and approaches the whole-report asymptote in about 200–300 ms. The performance difference between the partial- and whole-report conditions is called the *partial-report superiority effect*. The effect demonstrates performance benefits from one form of visual sensory memory, the iconic memory. The relationship between observer's performance and cue delay is called the *iconic memory decay function*.

Citation: Baek, J., Lesmes, L. A., & Lu, Z.-L. (2016). qPR: An adaptive partial-report procedure based on Bayesian inference. *Journal of Vision, 16*(10):25, 1–23, doi:10.1167/16.10.25.

doi: 10.1167/16.10.25 Received August 12, 2015; published August XX, 2016 ISSN 1534-7362

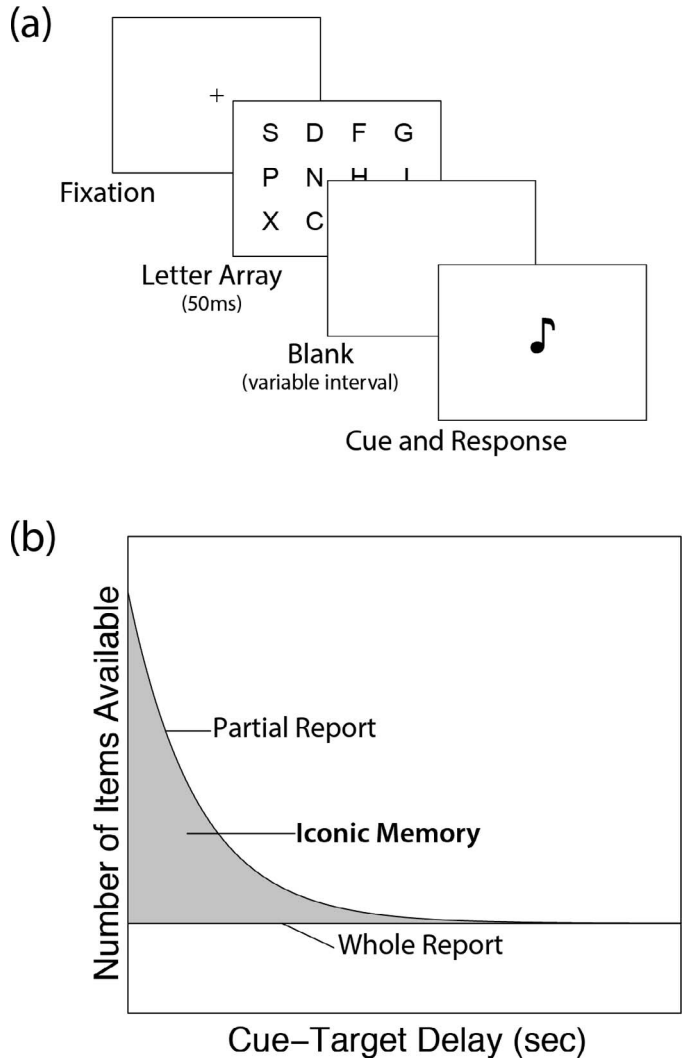

Figure 1. Partial-report experiment from Sperling (1960). (a) Illustration of a partial-report procedure. After a brief presentation of stimulus (a 3 × 4 array of letters) followed by a blank screen, observers were given a sound cue. They were asked to report letters of the cued row in the partial-report condition, and all letters in the whole report condition. (b) Exemplary results of a partial-report experiment. The performance difference between the partial-report and whole-report conditions (the shaded area) is called the partial-report superiority effect.

The partial-report paradigm has been accepted as the standard experimental paradigm to assess iconic memory, with several procedural modifications from the original version (Lu, Neuse, Madigan, & Dosher, 2005): (a) a visual cue, often in the form of an arrow in the center of the display, is used in place of the auditory cue (Averbach & Sperling, 1961); (b) letters are arranged on an imaginary circle instead of a rectangular array to equate eccentricity (Keele & Chase, 1967); (c) observers are asked to report the identity of a single cued letter instead of multiple letters (Averbach & Coriell, 1961); and (d) performance is scored as the probability of correctly reporting the cued letter instead of the total number of correctly reported letters (Averbach & Coriell, 1961).

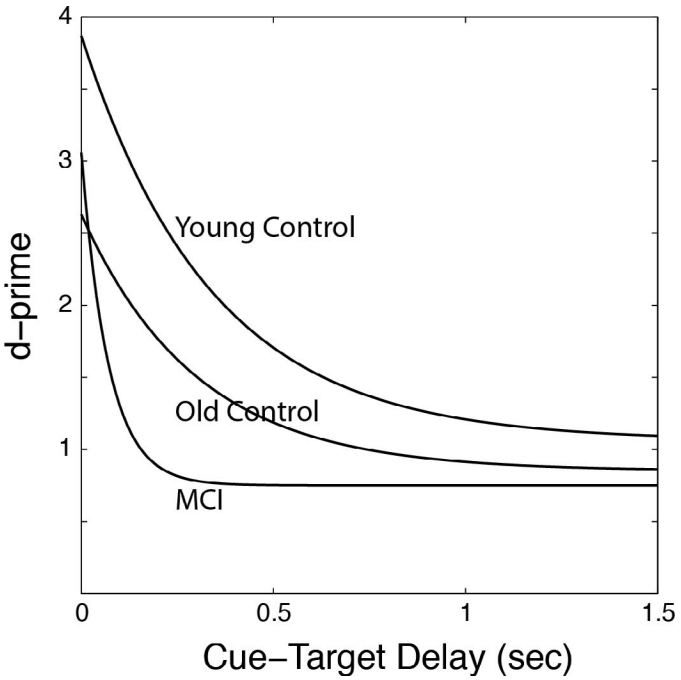

Figure 2. Iconic memory decay function of three groups of observers in Lu et al. (2005). Iconic memory decayed much faster in the MCI group than in the old control and young normal groups.

During the past half century, iconic memory has been extensively studied and widely accepted as a critical component in many theories of human information processing (Becker, Pashler, & Anstis, 2000; Coltheart, 1980; Dick, 1974; Di Lollo, 1980; Jonides, Irwin, & Yantis, 1982; Long, 1980). Recently, Lu et al. (2005) compared iconic memory decay functions of young normal, old normal, and old observers with mild cognitive impairment (MCI). They found that iconic memory decayed much faster in the MCI group than the others (Figure 2), although the two groups of older observers performed at equivalent levels in precue tests (assay of visibility) and tests cued at long delays (assay of short-term memory). Because more than 80% of people with MCI develop Alzheimer's disease (AD) in 10 years (Peripheral and Central Nervous System Drugs Advisory Committee, 2001), this finding suggests that fast decay of iconic memory might be an early sign of AD (Chong & Sahadevan, 2005; Dixon et al., 2007; Lu et al., 2005). In another study, Miller, Rammsayer, Schweizer, and Troche (2010) found that higher psychometric intelligence was associated with slower decay of iconic memory even when controlled for the influence of selective and divided attention. The relationship between iconic memory and intelligence

suggests preattentive process in iconic memory may play an important role in human information processing.

Conventionally, the method of constant stimuli (MCS) is used to measure the iconic memory decay function in partial report studies. Observer's response at a number of preselected cue delays is measured with about 100 trials per delay. An empirical iconic memory decay function is obtained by estimating the probability of correctly reporting the cued letter at each delay from the observer's responses. Often a theoretical curve such as an exponential decay function is fit to the empirical data to characterize the memory decay process. In a typical study, between 600 and 800 test trials (6–8 cue delays × 100 trials/delay) are necessary to obtain a good estimate of the iconic memory decay function. The measurement takes approximately one hour for normal young observers but much longer for observers in special populations (Lu et al., 2005). The long testing time makes it difficult or even impossible to carry out partial report experiments in special populations.

Many adaptive procedures have been developed to reduce the burden of data collection in psychophysical experiments (for reviews, see Leek, 2001; Lu & Dosher, 2013; Treutwein, 1995). Most developments have focused on characterizing the psychometric function, including various nonparametric procedures for estimating a single threshold (Derman, 1957; Kaernbach, 1991; Kesten, 1958; Levitt, 1971; Robbins & Monro, 1951; Taylor & Creelman, 1967; Tyrrell & Owens, 1988), and Bayesian adaptive procedures for estimating either a single threshold (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994; Watson & Pelli, 1983) or the threshold and slope of a psychometric function (Kontsevich & Tyler, 1999; Prins, 2013). Recent developments in this area have extended adaptive procedures to measure various psychological functions, including the threshold versus contrast function (Lesmes, Jeon, Lu, & Dosher, 2006), contrast sensitivity function (Dorr et al., 2015; Hou et al., 2010; Hou, Lesmes, Bex, Dorr, & Lu, 2015; Lesmes, Lu, Baek, & Albright, 2010) and sensitivity and bias parameters in Yes–No tasks (Kujala & Lukka, 2006; Kujala, Richardson, & Lyytinen, 2010; Lesmes et al., 2015). In all these adaptive procedures, the stimulus in the next trial is determined by the observer's previous responses to improve the efficiency of the test.

The goal of the present study is to alleviate the testing demand in the partial report procedure by developing a new adaptive procedure for estimating iconic memory decay function with a small number of trials without sacrificing accuracy and precision. We developed an adaptive partial report procedure, the quick Partial Report, or qPR, method, which is based on Bayesian inference and conducted a validation study of the new procedure.

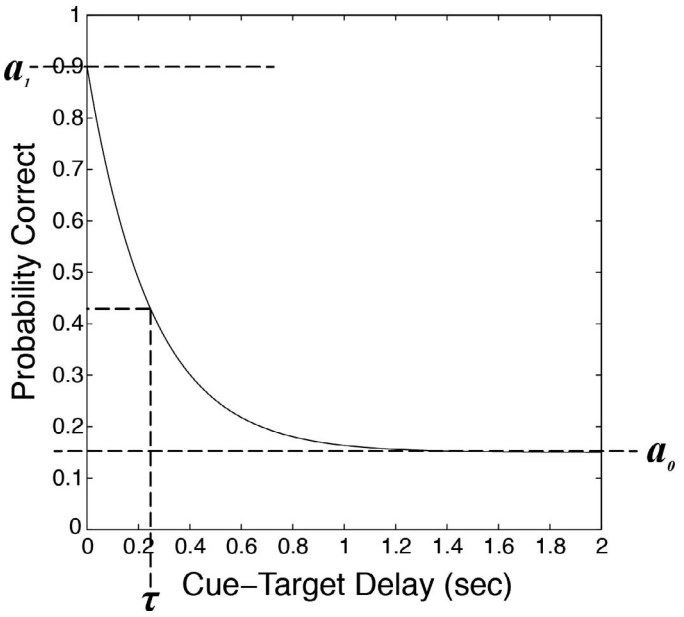

Figure 3. Parameterization of the iconic memory decay function. Temporal properties of iconic memory decay is characterized by an exponential decay function with three parameters: (a) a_0: performance at long delays (the amount of information transferred into short-term memory without the benefit of cuing); (b) a_1: performance with the simultaneous cue (initial visual availability of stimulus information); and (c)τ: time constant of iconic memory decay.

The quick partial-report method

The qPR method consists of four components. First, the iconic memory decay function is characterized with an exponential decay function and a broad prior distribution of the parameters. Second, following the observer's response in each trial, the posterior distribution of the parameters is updated using Bayes rule. Third, the stimulus for the next trial is selected to maximize the expected information gain on the parameters of the iconic memory decay function. Finally, the second and third steps are repeated until a fixed number of trials or a preset test precision is reached. Here we describe the qPR algorithm. The MATLAB simulation program is available for download (http://lobes.osu.edu/downloads/qPR_Publication_v1.0.1.zip).

Parameterization

The iconic memory decay function has been traditionally modeled with an exponential decay function (Lu et al., 2005; Figure 3):

$$pc(x) = a_0 + (a_1 - a_0)e^{-x/\tau}, \quad (1)$$

where $pc(x)$ is the probability of correctly reporting the

cued letter at a target and cue onset asynchrony of x; a_0 is the asymptotic performance level, often associated with residual information in short-term memory after iconic memory decays completely; a_1 is the performance level when $x = 0$; and τ is the time constant of iconic memory decay, the time it takes for the observer's partial-report superiority effect to drop to 37% of its initial level. In the beginning, the procedure defines (a) a three dimensional $\theta = (a_0, a_1, \tau)$ parameter space that represents all potential observable iconic memory decay functions, and (b) a one-dimensional stimulus search space over a range of possible levels of target and cue onset asynchronies, x.

Estimation: Bayesian update

The qPR procedure estimates three parameters of the iconic memory decay function using Bayesian inference. It starts with a prior probability distribution of the three parameters and updates their posterior probability distribution and the corresponding parameter estimates based on the observer's response after each trial. Bayesian inference is an important computational method in statistics, computer technology, and many domains of science (Androutsopoulos et al., 2000; Dawid, 2005; Kruschke, 2011; Yang & Rannala, 1997). It has been increasingly applied in psychophysics (Cobo-Lewis, 1997; Kontsevich & Tyler, 1999; Kujala & Lukka, 2006; Lesmes et al., 2006, 2010; Watson & Pelli, 1983) and clinical trials (Berry, Carlin, Lee, & Muller, 2010).

In the beginning of the qPR procedure, the prior probability distribution, $p_{t=0}(\theta)$, is defined as a three-dimensional joint probability distribution, in which each dimension corresponds to one parameter of the iconic memory decay function.

The prior represents a priori knowledge of the observer's iconic memory decay function. Often, it spreads over a broad range of parameter values. Alternatively, it can focus more narrowly on the likely values of the parameters if it can be informed by other characteristics of the observer (e.g., demographic information) and/or posterior distribution(s) of the same observer or other observers in previous study sessions (Kim, Pitt, Lu, Steyvers, & Myung, 2014). A more informative prior distribution could change the starting point of parameter estimation and make the estimation process faster. In the current implementation, the prior distribution is uniform over a wide range of parameter values.

After the t-th trial, the prior distribution $p_t(\theta)$ is updated to the posterior distribution $p_t(\theta \mid x, r_t)$ with the observer's response r_t (correct or incorrect) to a test with a cue delay of x by Bayes rule:

$$p_t(\theta|x, r_t) = \frac{p(r_t|\theta, x)}{p_t(r_t|x)} \times p_t(\theta), \tag{2}$$

where $\theta = (a_0, a_1, \tau)$ represents the parameters of the iconic memory decay function, $p_t(\theta)$ is the prior probability density function of θ. The probability of a response r_t in a given stimulus condition x, $p_t(r_t \mid x)$, is estimated by weighting the empirical response probability by the prior:

$$p_t(r_t|x) = \sum_{\theta} [p(r_t|\theta, x) p_t(\theta)], \tag{3}$$

where $p(r_t|\theta,x)$ is the likelihood of observing response r_t given θ and cue delay x; $p(r_t = \text{correct}|\theta,x)$ is computed with the iconic memory decay function (Equation 1); $p(r_t = \text{incorrect}|\theta,x) = 1 - p(r_t = \text{correct}|\theta,x)$.

The posterior $p(r_t|\theta,x)$ following the t-th trial serves as the prior $p_{t+1}(\theta)$ in the next trial:

$$p_{t+1}(\theta) = p_t(\theta|x, r_t). \tag{4}$$

Marginal posterior distributions of the parameters are computed via two-dimensional summation:

$$\begin{aligned} p_t(\vec{a}_0|x, r_t) &= \sum_{\tau} \sum_{a_1} p_t(\theta|x, r_t), \\ p_t(\vec{a}_1|x, r_t) &= \sum_{\tau} \sum_{a_0} p_t(\theta|x, r_t), \\ p_t(\vec{\tau}|x, r_t) &= \sum_{a_1} \sum_{a_0} p_t(\theta|x, r_t), \end{aligned} \tag{5}$$

and the means of the marginal posterior distributions are used to estimate the parameters of the iconic decay function after each trial:

$$\begin{aligned} \hat{a}_0 &= \sum [\vec{a}_0 \times p_t(\vec{a}_0)], \\ \hat{a}_1 &= \sum [\vec{a}_1 \times p_t(\vec{a}_1)], \\ \hat{\tau} &= \sum [\vec{\tau} \times p_t(\vec{\tau})]. \end{aligned} \tag{6}$$

In the current implementation of the qPR procedure, we define the prior on a grid and sum over all the points on the grid to compute parameter estimates. The effect of discrete grid has been investigated in other studies. One study showed that the precision of the estimated parameters with a discrete grid was comparable to MCMC-based continuous priors (Gu, 2012). The estimated iconic memory decay function can be computed with the estimated parameters using Equation 1. Alternatively, it can be estimated by a resampling method: the procedure resamples the posterior distribution of the qPR parameters 1,000 times, reconstructs the iconic memory decay functions from each set of sample parameters, and then averages over all the resampled iconic memory decay functions. In this implementation, we used the latter method. This resampling procedure automatically takes into account of the covariance structure in the posterior distribution of the qPR parameters, and can be used to directly assess the variability of the estimated iconic memory decay function.

Stimulus selection: One-step ahead search for minimum entropy

To select the stimulus for the next trial, the qPR determines the cue delay x that would maximize the expected information gain about the parameters of the iconic memory decay function. Here, information is quantified by entropy, a measure of uncertainty associated with variables (Shannon, 1948). The qPR uses a one-step-ahead search strategy to search for the cue delay that would lead to the minimum expected entropy (Cobo-Lewis, 1997; Kontsevich & Tyler, 1999; Kujala & Lukka, 2006; Lesmes, et al., 2006). It first computes the expected posterior probability distributions for all possible cue delays. The entropy of the posterior is defined as:

$$H_{t+1}(x,r) = -\sum_{\theta} p_{t+1}(\theta|x,r)\log\Big(p_{t+1}(\theta|x,r)\Big) \quad (7)$$

The qPR then computes the observer's response probability $p_{t+1}(r|x)$ in every possible cue delay condition in the next trial based on the current prior. The expected entropy after a trial with cue delay x is calculated as a weighted sum of posterior entropy:

$$E[H_{t+1}(x,r)] = \sum_{r} H_{t+1}(x,r)p_{t+1}(r|x) \quad (8)$$

The cue delay with the minimum expected entropy is chosen for the next trial:

$$x_{t+1} = \arg\min_{x} E[H_{t+1}(x)]. \quad (9)$$

This is equivalent to maximizing the expected information gain, quantified as the entropy change between the prior and posterior (Kujala & Lukka, 2006; Lesmes et al., 2006). Figure 4 shows an example of stimulus selection and information gain in a qPR experiment. In the beginning of the procedure, the expected information gain after the first trial is maximum at $x = 0.11$ s (bottom left panel) so that a cue delay of 0.11 s is tested in the first trial (top panel). After the 10th trial, qPR expects that the information gain after the 11th trial would be maximum at $x = 0$ s. The procedure therefore selects the stimulus with a cue delay of 0 s in the 11th trial.

Stopping rule

The qPR procedure terminates after a fixed number of trials. Alternatively, the qPR procedure can stop after it achieves a certain defined objective (e.g., after reaching a criterion level of precision for all the parameters, the decay function, or the decay time constant). Although the dynamic stopping rule has been reported to be more efficient than the fixed-length procedure, its superiority occurs in experiments with short runs (Alcalá-Quintana & García-Pérez, 2005; Tanner, 2008). In our validation study, 200 qPR trials were used, but simulation results showed that only 100 trials were sufficient to estimate the iconic memory decay function with reasonable precision.

Simulation

Method

To evaluate the performance of the qPR procedure for observers with a range of iconic memory performance, we simulated three observers with distinct parameter values that approximately correspond to the young, old normal, and old MCI groups in Lu et al. (2005). The parameters of the simulated observers are summarized in Table 1.

The range of parameters was set based on results of previous iconic memory studies (Lu et al., 2005; Sperling, 1960, 1967), with 41 linearly spaced a_0 values (from 0 to 0.5), 41 linearly spaced a_1 values (from 0.5 to 1.0), and 40 log-linearly spaced τ values (from 0.01 to 0.8 s). The broad parameter space ensures robust assessment of various populations and avoids effects of extreme values—the tendency to bias toward the center of the parameter space when the observer's true parameter values are close to the boundary of the space. The prior was set to a uniform distribution in the current study.

Possible cue delays were sampled from 0 to 3 s with 30 equally spaced samples on a logarithmic scale. The range was broader than the range of 0 to 1.5 s commonly used in partial report studies. Including test trials at long cue delays is important for accurate estimation of the asymptotic level a_0 of the iconic memory decay function.

The simulated observers performed a 10-alternative forced-choice (10AFC) letter identification task in a partial report procedure. Ten letters, each drawn randomly from a set of 10 letters and arranged on an imaginary circle, were presented to the observers in each trial. The observers must identify the letter in the cued location in each trial. Each simulated experimental run consisted of 1,600 trials. In each trial, the expected probability correct, $pc(x)$, of the simulated observer was calculated for the selected cue delay x (Equation 1). Observer's response in each trial was simulated by drawing a random number r from a uniform distribution over the interval from 0 to 1. The response was labeled as correct if $r < pc(x)$, and incorrect otherwise. 1,000 simulated runs were conducted for each observer.

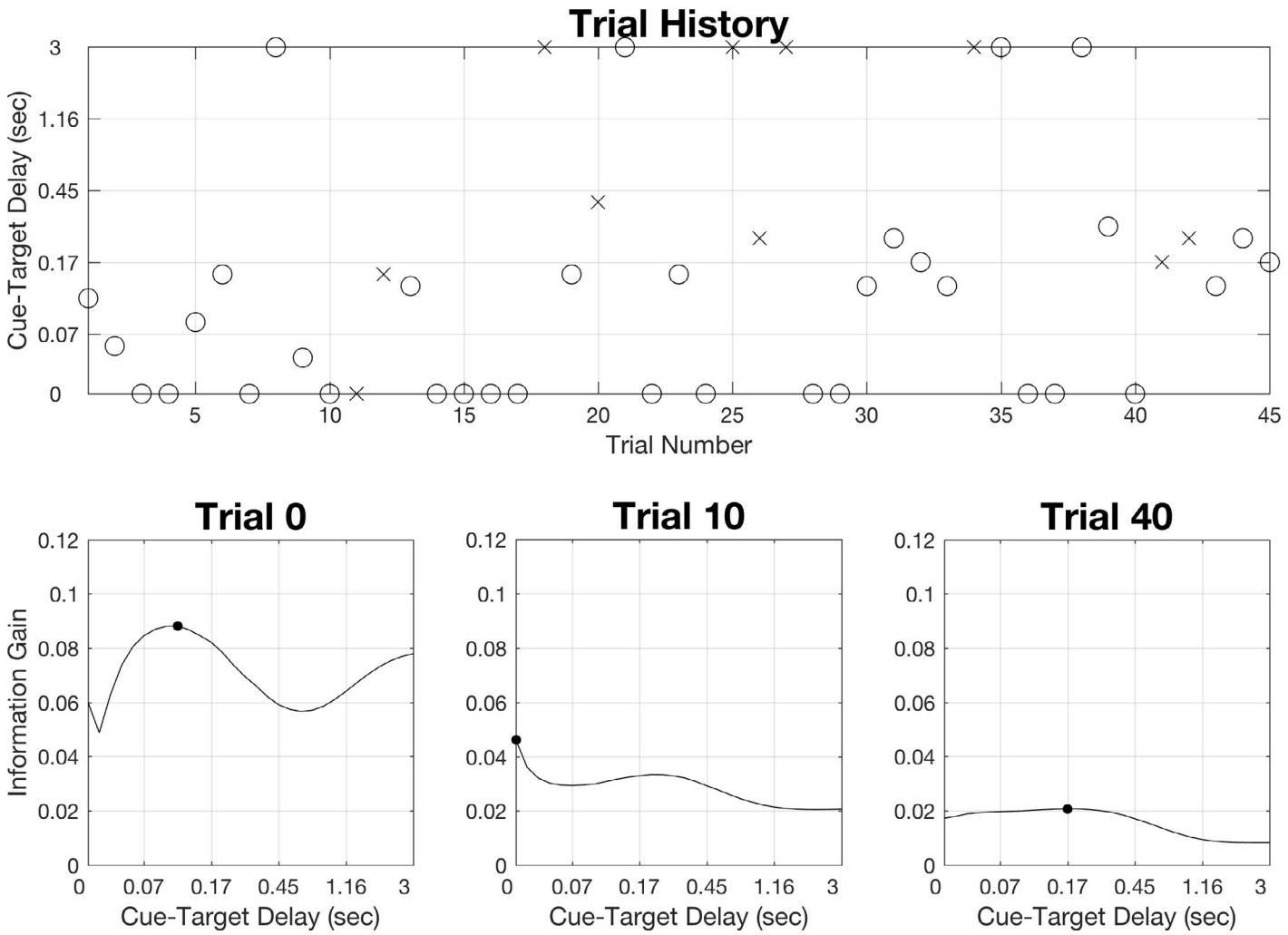

Figure 4. An example of stimulus selection and information gain. Top panel: Tested cue delays as a function of trial number. Circles represent correct responses and crosses represent incorrect responses. Bottom panels: Expected information gain in trial 1 (before running the procedure), 11, and 41, computed after finishing 0, 10, and 40 trials. The cue delay with maximum information gain (filled circles in the bottom panels) is tested in the next trial. Before the procedure, for example, the expected information gain after the first trial is maximum at cue delay = 0.11 s (bottom left panel) so that the stimulus with cue delay = 0.11 s is tested in the first trial (top panel).

Evaluation: Accuracy and precision

Accuracy is a measure of how much the estimates deviate from the truth on average, and precision is a measure of the variability of the estimates (Lu & Dosher, 2013; Treutwein, 1995). A good procedure should quickly increase the accuracy of the estimated iconic memory decay function or its parameters as trial number increases and lead to an unbiased estimate.

	Observer 1 (young normal)	Observer 2 (old normal)	Observer 3 (MCI)
a_0	0.35	0.25	0.25
a_1	0.95	0.85	0.85
$\tau(s)$	0.30	0.30	0.05

Table 1. Parameters of the three simulated observers.

Bias can be calculated by the mean discrepancy between the estimated and true parameters of the iconic memory decay function. For each parameter of the iconic memory decay function, the bias of the estimate after i-th trial can be calculated as:

$$\text{bias}_i = \frac{\sum_{j=1}^{J}(P_{ij} - P_{\text{true}})}{J}, \quad (10)$$

where P_{true} is the true parameter value, and P_{ij} is the parameter estimate obtained after the i-th trial in the j-th simulation.

Because the empirical data in partial report experiments is the iconic memory decay function (i.e., probability correct as a function of cue delay) rather than its parameters, it is also important to assess the

performance of qPR in terms of its estimate of the iconic memory decay function (i.e., probability of correctly reporting the cued letter in a range of cue delays). The average absolute bias of the estimated iconic decay function after the i-th trial can be calculated as[1]:

$$\text{average absolute bias}_i = \frac{\sum_k \left| \sum_j (pc_{ijk} - pc_k^{\text{true}}) \right|}{J \times K}, \quad (11)$$

where pc_{ijk} is the estimated probability correct in the k-th cue delay condition after i trials obtained in the j-th simulation and pc_k^{true} is the true probability correct.

Precision is defined as the inverse of the variability of the estimates. Two methods have been used to assess the precision of the qPR procedure. The first is based on the standard deviation of repeated measures:

$$\text{SD}_i = \sqrt{\frac{\sum_k \sum_j \left(pc_{ijk} - \text{mean}(pc_{ijk}) \right)^2}{J \times K}} \quad (12)$$

Another measure of precision is the half width of the credible interval (HWCI) of the posterior distribution of the estimated iconic memory decay function (Edwards, Lindman, & Savage, 1963). The HWCI refers to the shortest interval that covers most of the distribution (see Figure 5 for an example). The 95% credible interval represents the range within which the actual value lies with 95% probability (Clayton & Hills, 1993), whereas confidence interval, the most popular index of precision, represents an interval that contains the true value of the parameter for 95% of unlimited repetitions (Rothman & Greenland, 1998). Since researchers typically do not iterate an experiment many times for the same observer, the HWCI of the posterior distribution is a good index of performance that can be obtained with a single run of the qPR procedure. This is one major advantage of the Bayesian inference framework.

Results

Figure 6 shows the history of qPR parameter estimates for the three simulated observers over 1,600 trials. The bias of the estimated parameters rapidly decreased toward zero during the first 20–50 trials and the precision of the estimated parameters increased with trial number. For a_0, the average bias became less than 0.017 after 100 trials and 0.010 after 200 trials, and further decreased to 0.005 after 400 trials, 0.003 after 800 trials, and 0.001 after 1,600 trials. The average bias for a_1 was −0.019 after 100 trials and −0.011 after 200 trials, and further decreased to −0.005 after 400 trials, −0.003 after 800 trials, and −0.001 after 1,600 trials. For τ, the average bias was −0.013 after 100 trials and −0.003 after 200 trials, and decreased further to −0.0014 after 400 trials, −0.0015 after 800 trials, and −0.001 after 1,600 trials, all in log10 units. The average 68.2% HWCI of a_0 started at 0.157 in the first trial, and decreased to 0.071 after 100 trials, 0.054 after 200 trials, and 0.020 after 1,600 trials. The average 68.2% HWCI of a_1 started at 0.160 in the first trial, and decreased to 0.050 after 100 trials, 0.038 after 200 trials, and 0.014 after 1,600 trials. The average 68.2% HWCI of τ started at 0.526 in the first trial, and decreased to 0.216 after 100 trials, 0.154 after 200 trials, and 0.051 after 1,600 trials, all in log10 units. Since the accuracy and precision after qPR 200 trials were reasonably good, we only present results from the first 200 trials in the rest of this article. The accuracy and precision of estimates of the parameters of the iconic memory decay function after 20, 50, 100, and 200 trials are summarized in Table 2.

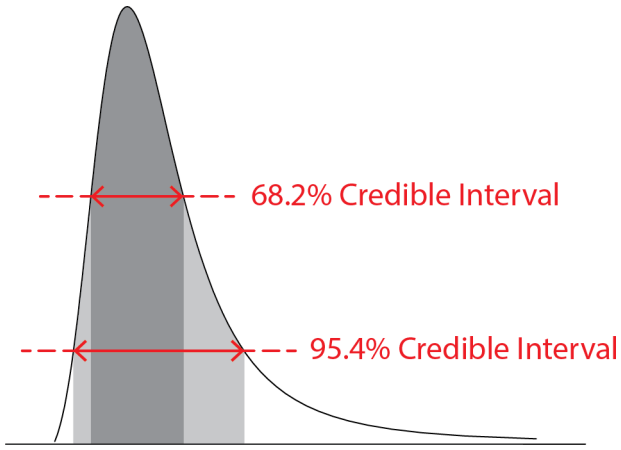

Figure 5. Bayesian credible interval. Shaded regions with dark gray and light gray represent 68.2% and 95.4% credible intervals of the probability distribution, respectively. A credible interval represents the shortest interval that covers a particular percentage (e.g., 68.2% or 95.4%) of total area. In contrast to confidence interval, the lower and upper bounds always have the same probability densities and the probability densities within the credible interval are greater than those outside of the interval, even in asymmetric distributions.

Figure 7a shows the accuracy and precision of the estimated iconic memory decay functions obtained with 20, 50, 100, and 200 qPR trials. Results of the three simulated observers are presented in different columns. The true iconic memory decay functions are plotted as dashed curves, and the qPR estimates are shown as solid curves. Shaded areas represent the 68.2% HWCI of the qPR estimates. With increasing trial numbers, the qPR improved both the accuracy (decreasing discrepancy between dashed and solid lines) and

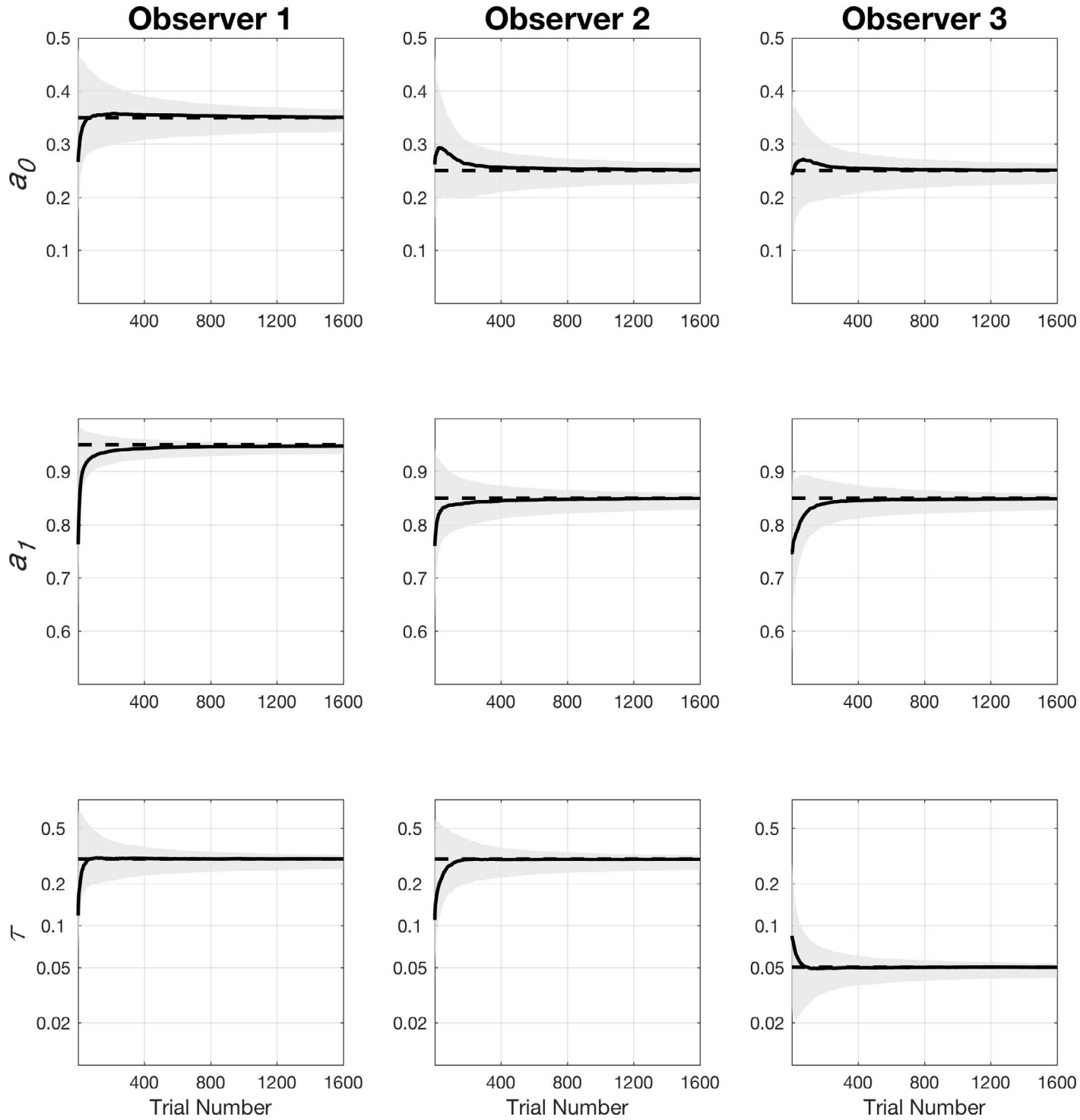

Figure 6. Simulation results. Results on three simulated observers with distinct parameters are presented in different columns. The parameters of the three observers correspond approximately to the young normal, old normal, and MCI groups in Lu et al. (2005). Each observer's parameter estimates for a_0 (top), a_1 (middle), and τ (bottom) are plotted as a function of trial numbers. Dashed lines represent true parameter values and solid lines represent the mean of the estimated parameters. Shaded regions denote 68.2% HWCIs. The estimates rapidly converged to the true values (50–100 trials).

precision (decreasing shaded area) of the estimated iconic memory decay functions. As trial number increased (50–200 trials), the discrepancy between the true (dashed lines) and estimated (solid lines) decay functions decreased. It took fewer than 100 trials to recover the general shape of the iconic memory decay function (average absolute bias = 0.026). The average 68.2% HWCI became less than 0.1 after 36 trials for all

	a_0	a_1	τ (log10 unit)
Observer 1			
20 trials	−0.027 (0.107)	−0.062 (0.065)	−0.109 (0.320)
50 trials	−0.004 (0.086)	−0.034 (0.045)	−0.023 (0.242)
100 trials	0.003 (0.071)	−0.021 (0.034)	0.005 (0.192)
200 trials	0.007 (0.056)	−0.012 (0.026)	0.003 (0.145)
Observer 2			
20 trials	0.041 (0.109)	−0.036 (0.092)	−0.213 (0.371)
50 trials	0.040 (0.089)	−0.019 (0.071)	−0.114 (0.284)
100 trials	0.029 (0.072)	−0.013 (0.057)	−0.043 (0.218)
200 trials	0.013 (0.053)	−0.010 (0.043)	−0.005 (0.157)
Observer 3			
20 trials	0.010 (0.109)	−0.072 (0.102)	0.116 (0.426)
50 trials	0.019 (0.089)	−0.046 (0.078)	0.037 (0.325)
100 trials	0.019 (0.071)	−0.024 (0.059)	−0.002 (0.237)
200 trials	0.010 (0.053)	−0.010 (0.043)	−0.007 (0.159)

Table 2. Accuracy and precision of the qPR estimates of the parameters of the iconic memory decay function for three simulated observers. The 68.2% HWCIs are given in parentheses.

observers, and decreased to below 0.075 after 86 trials and 0.05 after 214 trials. The accuracy (average absolute bias) and precision (68.2% HWCI) of the estimated iconic memory decay function are summarized in Figure 7b and Table 3. The results indicate that the qPR can rapidly estimate the true iconic memory decay function for a wide range of potential populations with only a small number of trials.

Comparison between the qPR and MCS

To directly compare the efficiency of the qPR and MCS, we computed the expected standard deviation of the iconic memory decay function estimated with the MCS procedure as a function of number of test trials by:

$$SD_i^{MCS} = \frac{\sum_i \sqrt{\frac{pc_i(1-pc_i)}{n_i}}}{I} \quad (13)$$

where pc_i is the true probability correct at the i-th cue delay, n_i is the number of trials tested at the i-th cue delay. Here we assume that eight cue delays ($I = 8$) are tested with the MCS method so that the total number of tested trials is $n_i \times 8$. The eight cue delays, 0, 0.03, 0.06, 0.14, 0.30, 0.65, 1.4, and 3.0 s, are selected based on a typical iconic memory procedure (Lu et al., 2005) and are the same as those used in the psychophysical validation experiment. A comparison between the standard deviation of the estimates from the MCS, the standard deviation of the estimates from the qPR, and the average 68.2% HWCI of the qPR is presented in Figure 8 and Table 4. To reach a 0.075 probability correct precision, the MCS requires 240–270 trials, while the qPR only needs about 50–90 trials. The qPR is 3–5 times more efficient than the MCS. In terms of testing time, using the qPR, an iconic memory decay function can be estimated with reasonable accuracy and precision in less than 10 min, which is considerably faster than the 1 hr of testing time with the conventional MCS.

Comparison of precision indexes

In Figure 8, we present different indices of qPR precision: 68.2% HWCI and standard deviation of the qPR estimates. The difference between the two metrics is less than 0.01 after 100 trials (0.007, 0.002, and 0.000 for the three observers, respectively). The smaller estimated standard deviation in the beginning of the procedure is due to the fact that the qPR always starts with the same uniform prior distribution in repeated measures.

Sampling and information gain

Stimulus sampling pattern of the qPR procedure for observer 1 is presented in Figure 9. The diameter of each circle represents the proportion of trials tested at the cue delay represented by the x-coordinate of the center of the circle in trials 1–20, 21–50, 51–100, and 101–200. In the first 20 trials, the qPR intensively tests the shortest and longest cue delays to characterize a_0 and a_1, with only a few trials on short cue delays (0.03–0.3 s). Then the stimulus sampling of the qPR spreads to short cue delays and progressively moves to the range around the true τ value (0.2–0.4 s). The method does not frequently test cue delays greater than 0.5 s (except 3.0 s) throughout the whole experiment.

Figure 10 shows the trial-by-trial and cumulative information gain of the qPR and MCS procedures for the three simulated observers as functions of number of trials. It is evident that the qPR gains information faster than the MCS procedure. In both procedures, earlier trials provide much more information than later trials. For example, for simulated observer 1, the information gain of the first trial is 6.75 times that of the 100th trial, and 11.9 times that of the 200th trial in the qPR procedure. After 24, 48, and 96 trials, the cumulative information gain of the qPR procedure is 1.29, 1.34, and 1.41 times of that of the MCS procedure for simulated observer 1.

Effects of the prior

In these simulations, the prior probability distribution was set to a uniform distribution over all dimensions of the parameter space. Alternatively, it is possible to use prior knowledge to focus more narrowly on likely values of the parameters. It is well known that a different setting of the initial prior distribution would change the starting point of parameter estimation and

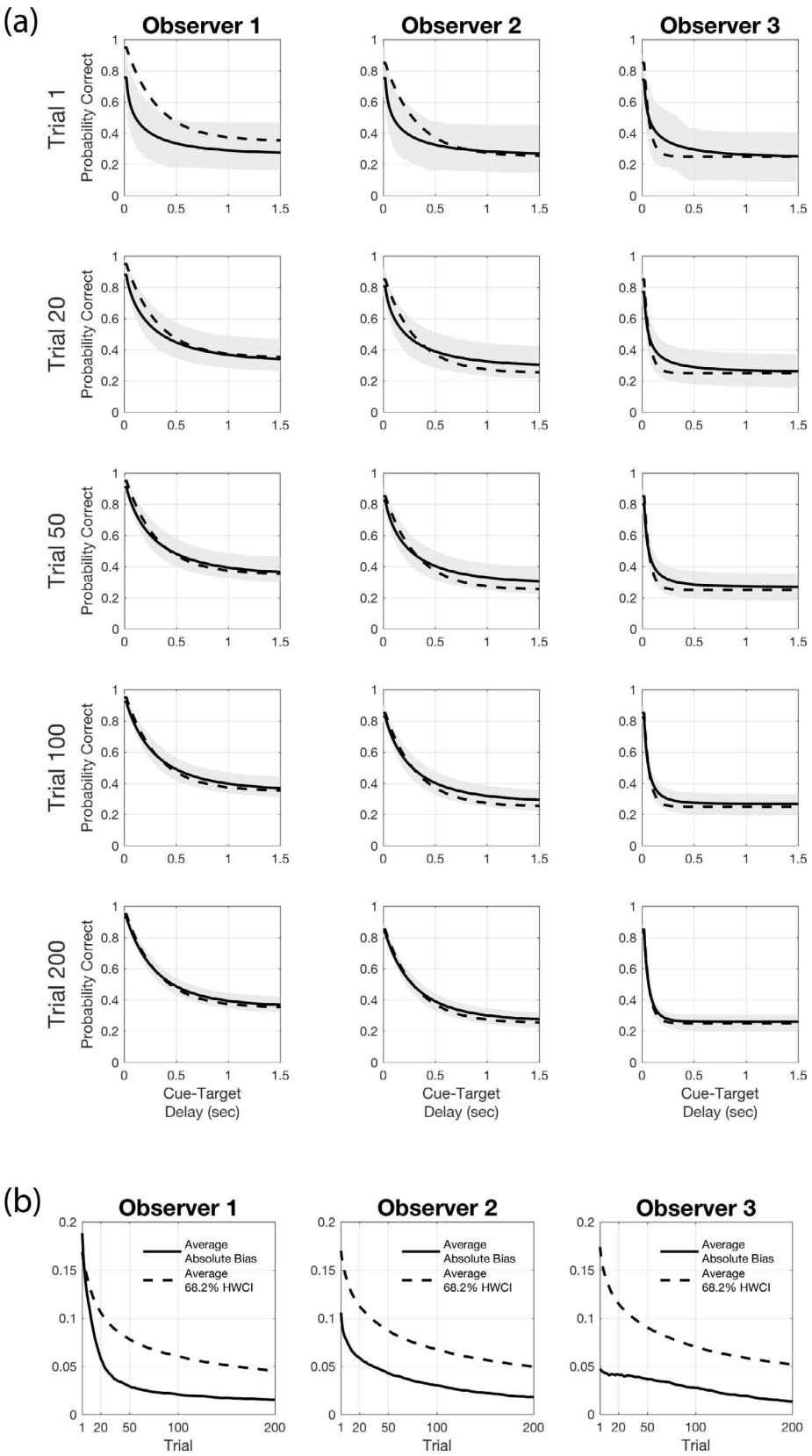

Figure 7. Simulation results. (a) Results of the three simulated observers are presented in different columns. Estimates obtained with 20, 50, 100, and 200 qPR trials are presented in different rows. The true iconic memory decay functions are plotted as dashed lines; the average qPR estimates are shown as solid lines. Shaded areas represent the 68.2% credible interval of qPR estimates. With increasing trial numbers, qPR improves accuracy (decreasing discrepancy between dashed and solid curves) and precision (decreasing shaded area). (b) Average absolute bias and precision of the estimated iconic memory decay function are plotted as a function of trial numbers. Solid lines represent average absolute bias (Equation 11); dashed lines represent 68.2% HWCI.

	Observer 1	Observer 2	Observer 3
20 trials	0.059 (0.106)	0.059 (0.113)	0.041 (0.115)
50 trials	0.029 (0.078)	0.043 (0.087)	0.037 (0.090)
100 trials	0.021 (0.060)	0.030 (0.067)	0.027 (0.070)
200 trials	0.015 (0.045)	0.018 (0.050)	0.013 (0.052)

Table 3. Average absolute bias and precision of the qPR estimate of the iconic memory decay function for three simulated observers. The average 68.2% HWCIs are given in parentheses (units: probability correct).

make the estimation process even more efficient (Gu et al., 2016). To illustrate the effects of the prior, another set of simulations was conducted with five different prior settings: (a) a uniform prior distribution, (b) a weakly informative proper prior, (c) a weakly informative but improper prior, (d) an informative proper prior, and (e) an informative but improper prior. Figure 11 shows that the qPR with a weakly informative prior—either proper or improper—has essentially the same performance as the one with a uniform prior after 30 trials, and that an informative proper prior can enhance the performance of the procedure, but there is a risk of getting deteriorated accuracy when the prior is informative but improper. In practice, the prior for the qPR procedure can be informed by prior knowledge or pilot data, such as the representative parameter distribution of a known population obtained before testing a particular individual.

Effects of lapse

One common problem in Bayesian adaptive procedures is that they are often vulnerable to mistakes (e.g., finger errors) made at the beginning of an experiment (Kontsevich & Tyler, 1999). Because the qPR uses sequential Bayesian inference, observer mistakes in the first few trials might mislead the procedure and cause inaccurate and imprecise estimates. To investigate the effect of lapse, observer 1 was simulated with random responses in the first 1, 3, and 5 trials. In the lapse trials, the expected probability correct was at the chance performance level for all cue delays. Figure 12 shows that a small number of lapse trials in the beginning of the experiment does not have much impact on precision, but significantly slows down accurate estimation. To reach a less than 0.05 average absolute bias, for example, 12, 19, 30, and 45 trials are required with lapse in the first 0, 1, 3, and 5 trials, respectively. However, the qPR recovers from observer's initial lapse and obtains sufficiently accurate estimation (average absolute bias $<$ 0.02) after 100–130 trials in all the simulated conditions.

Target precision	Methods	Observer 1	Observer 2	Observer 3
0.075	MCS	240	272	272
	qPR, *SD*	–	57	71
	qPR, 68.2% HWCI	56	77	86

Table 4. Number of trials required to reach a target precision (in probability correct).

Psychophysical validation

In addition to the simulation study, the performance of the qPR method was evaluated in a psychophysical

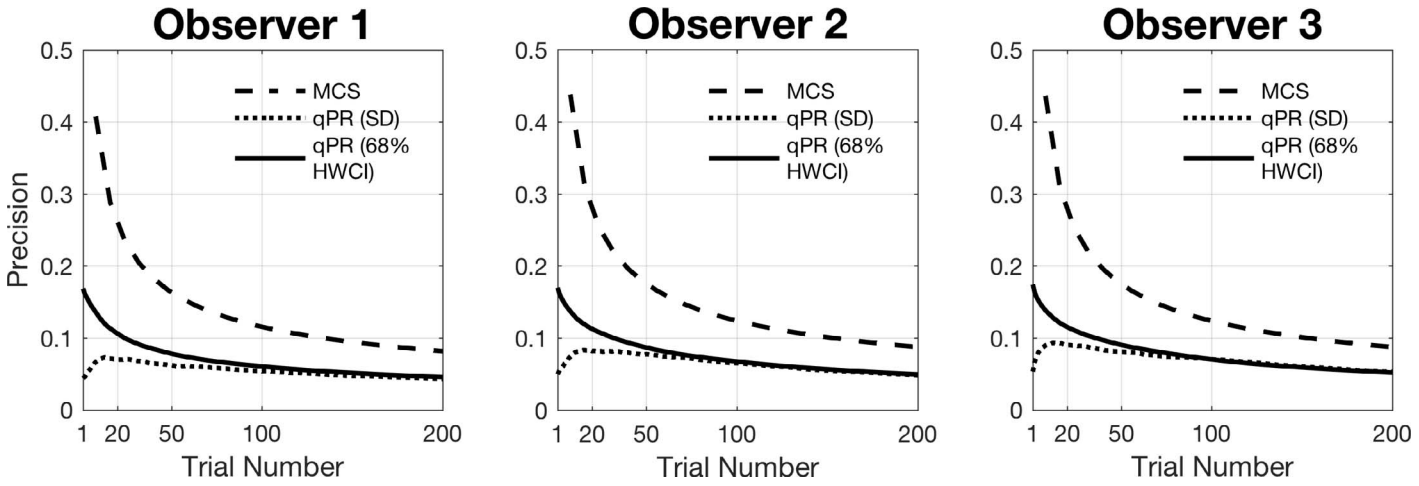

Figure 8. Comparison of the precision of the qPR and MCS procedures for the three simulated observers. The comparison was made between the standard deviation of the MCS estimates (dashed lines), the standard deviation of qPR estimates (dotted lines), and the average 68.2% HWCI of qPR estimates (solid lines). Note that the standard deviation and HWCI curves of the qPR estimates virtually overlap with each other after about 100 trials. The difference between the two metrics is less than 0.01 after 100 trials (0.007, 0.002 and 0.000 for the three observers, respectively).

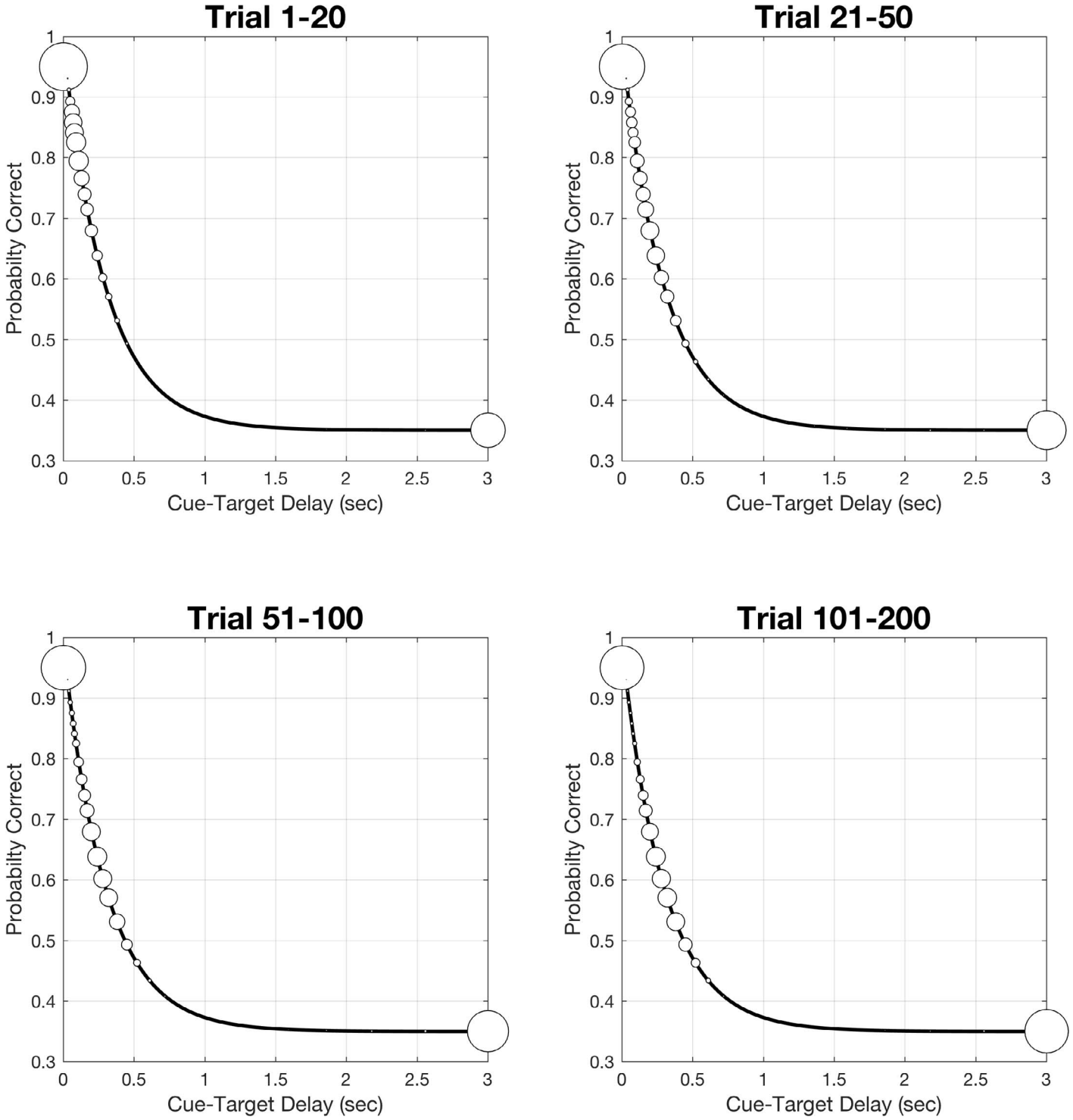

Figure 9. Stimulus sampling for simulated observer 1. Relative frequency of tested trials is presented as a function of cue delay, overlaid on the true iconic memory decay function. The diameter of the circles represents the proportion of tested trials in trials 1–20, 21–50, 50–100, and 101–200 trials in each cue delay condition over 300 simulations. At the beginning, sampling is focused on the shortest (0 s) and longest (3 s) cue delays to specify a_0 and a_1. Then it spreads to small cue delays (0–0.3 s) and moved to the true τ value (0.3 s).

experiment. We found that the iconic memory decay functions estimated with 100 qPR trials had good precision and excellent agreement with those obtained with 1,600 MCS trials.

Method

Two naive observers (FH and TC) and one of the authors (JB) participated in the experiment. All

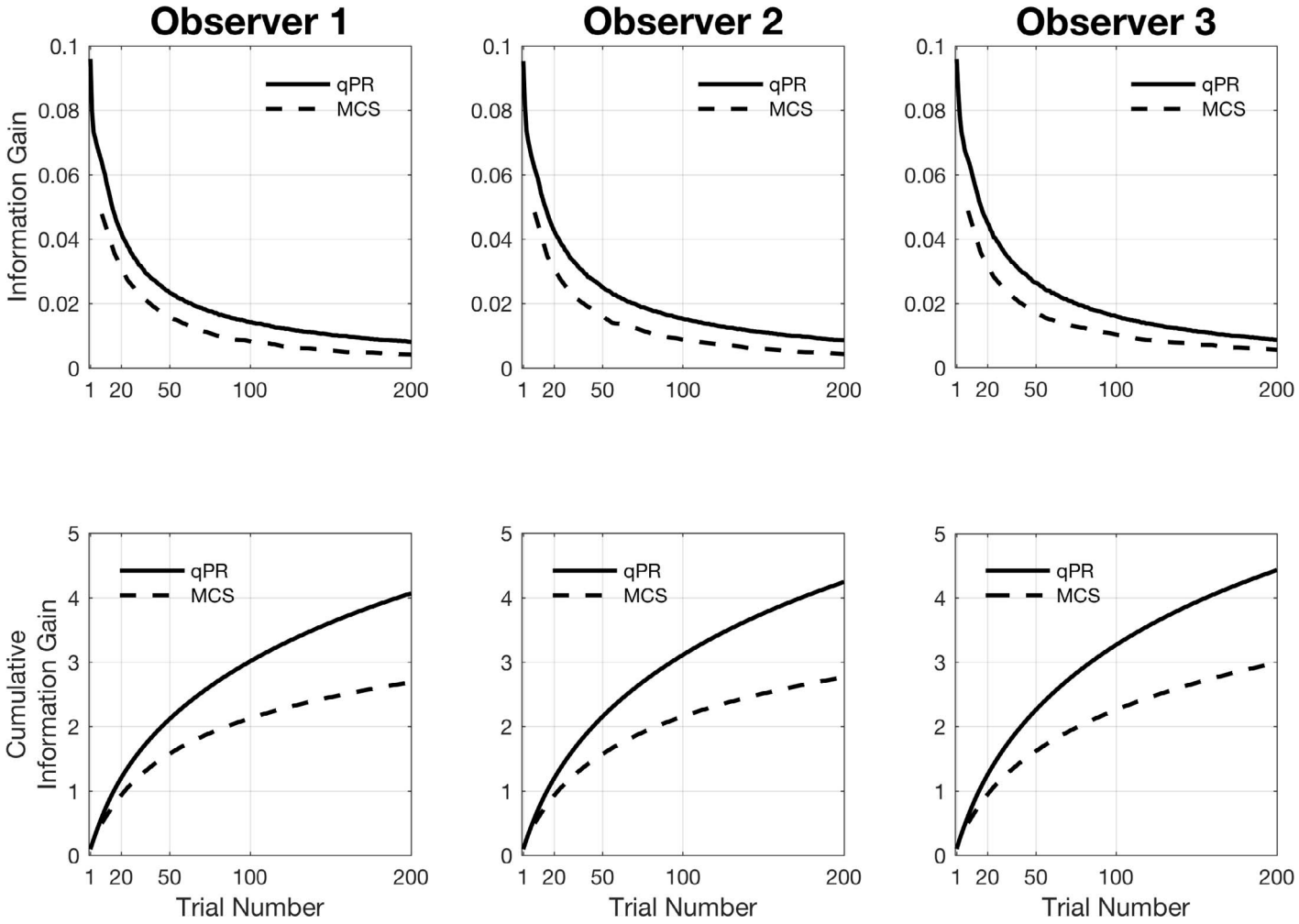

Figure 10. Trial-by-trial (upper row) and cumulative (bottom row) information gain for the three simulated observers. The information gain with the qPR and MCS procedures are presented as functions of trial number.

observers were males between 34 and 44 years of age and had corrected-to-normal vision. They were experienced in psychophysical studies.

The experiment was carried out on an IBM PC compatible computer. The stimuli were displayed on a Dell 17-in. color CRT monitor with a refresh rate of 100 Hz and viewed binocularly with natural pupil at a viewing distance of approximately 85 cm in a dimly lighted room. MATLAB programs written with Psychtoolbox extensions (Brainard, 1997; Pelli, 1997) were used to control the display and collect observers' responses. The qPR algorithm and the MCS were used to select stimulus conditions and measure the iconic memory decay function.

Eight qPR and one MCS measures were collected for each observer in four testing sessions. Each session consisted of 400 qPR trials and 400 MCS trials: two interleaved qPR runs with 200 trials each, and 50 MCS trials at each of eight cue delay conditions. All trials were randomly mixed in each session.

Each trial began with a fixation cross in the center of the display, followed by a 20 ms stimulus display containing ten letters. The letters were randomly selected from C, D, H, K, N, O, R, S, V, and Z without repetition, and placed with equal spacing on an imaginary circle with an eccentricity of 3.5° in a random order. Each letter subtended a visual angle of $1.25° \times 1.25°$ and was drawn in capital Sloan optotype to control perceptual legibility. The letter display was followed by a central cue pointing to one of the letters, with a selected cue delay. For the MCS trials, eight cue delays were used: the arrow cue was presented 0, 0.03, 0.06, 0.14, 0.30, 0.65, 1.4, or 3.0 s after the onset of the letter display. For the qPR trials, the procedure selected cue delays among 30 log-linearly spaced values between 0 and 3 s. In both types of trials, the cue remained on the screen till response. Observers were asked to report the cued letter using the computer keyboard, and a beep followed each incorrect response. An example of the trial sequence is

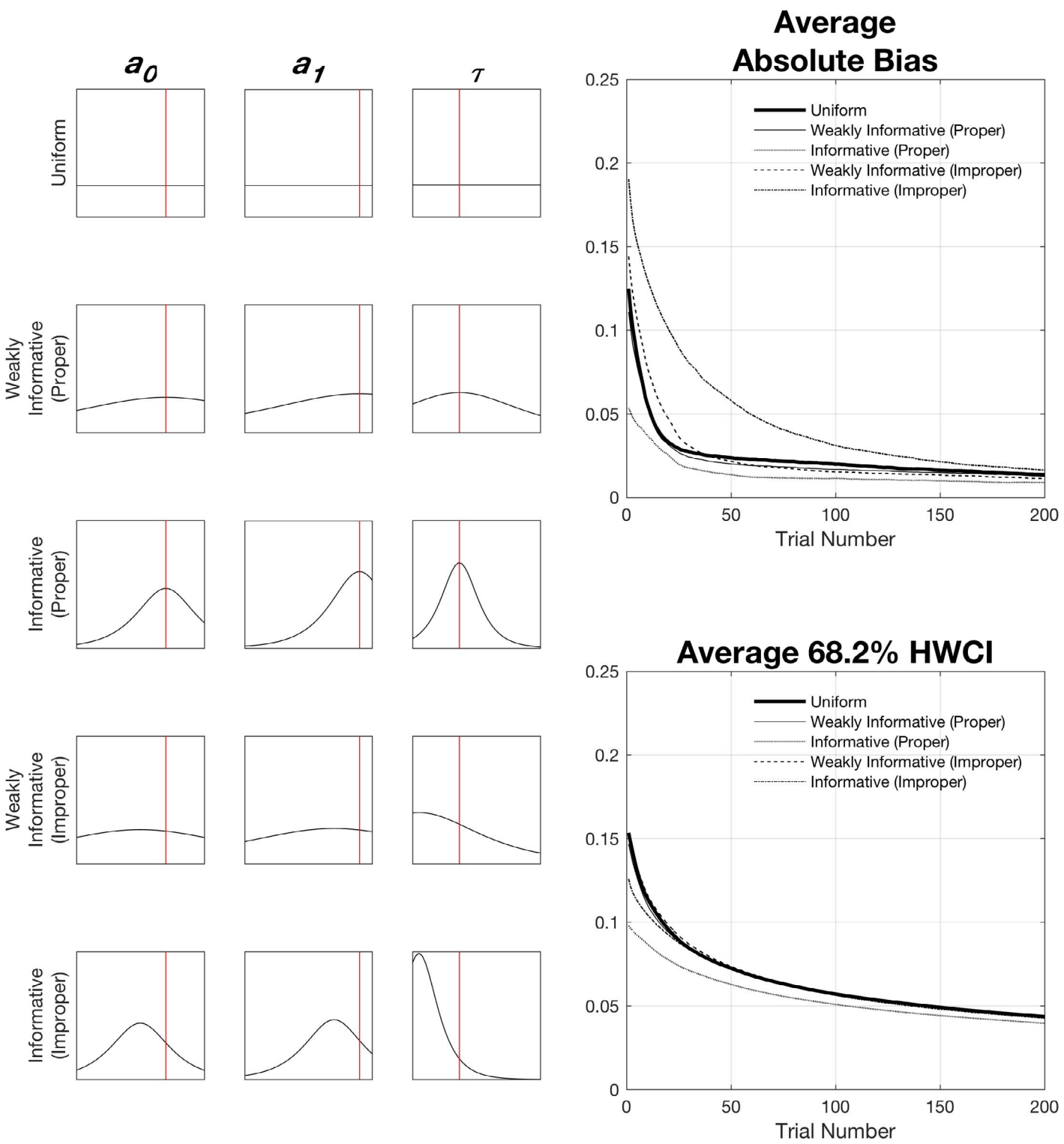

Figure 11. Effects of the prior for simulated observer 1. Figures in the left panel show five different prior settings. The performance of the qPR is shown in the right panels. The qPR with weakly informative priors—either proper or improper—has essentially the same performance as the one with a uniform prior after 30 trials and that an informative proper prior can enhance the performance of the procedure, but there is a risk of getting deteriorated accuracy when the informative prior is improper.

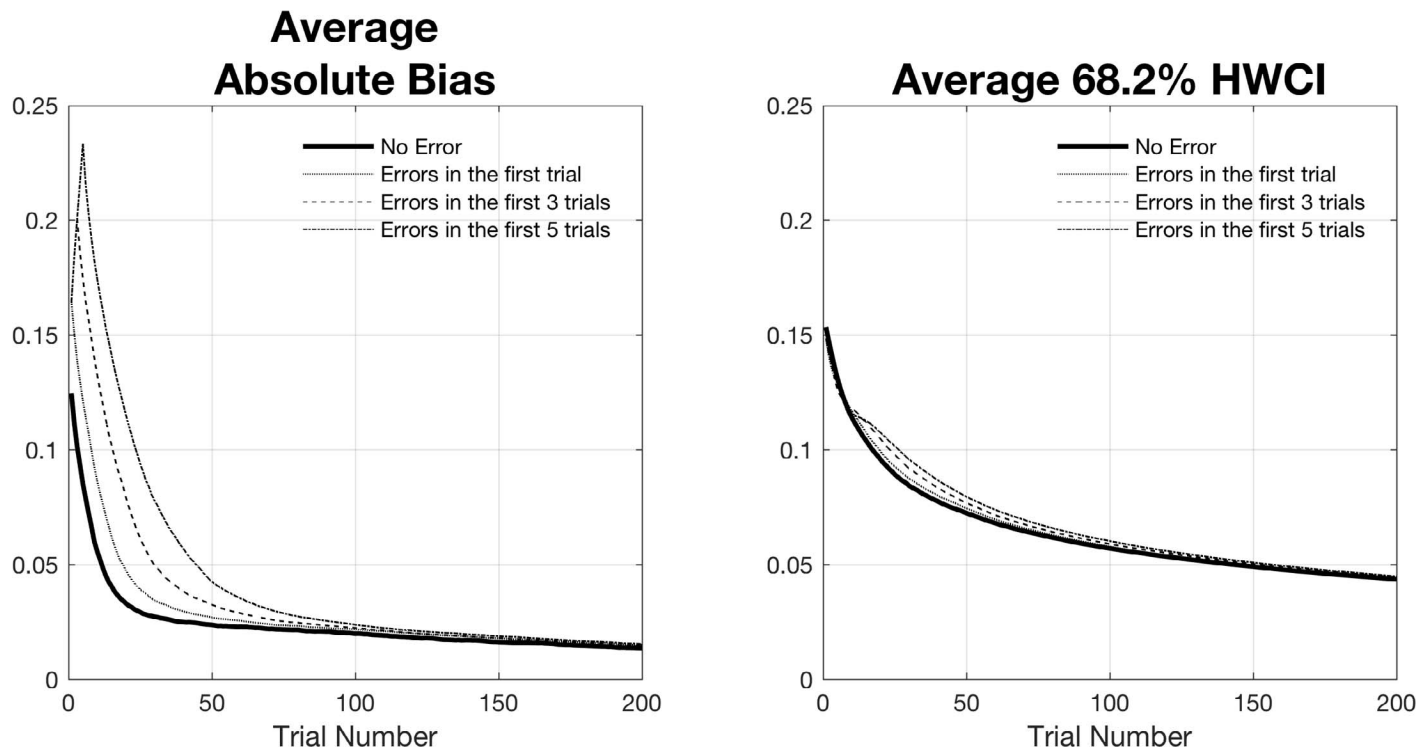

Figure 12. Effects of lapse in the first few trials. Observer 1 was simulated with random responses in the first 0, 1, 3, and 5 trials. A few mistakes significantly increased the average absolute bias of the estimates but did not have much impact on precision.

illustrated in Figure 13. Each session took approximately one hour, with a short break after every 200 trials.

Results

Figure 14 presents iconic memory decay functions measured with eight repeated runs of the qPR procedure (solid lines) and one run of 1,600 MCS trials (filled circles). The data from the MCS method was fitted with an exponential decay function (Equation 1) using a maximum likelihood procedure (dashed lines). The error region (shaded area) represents the average 68.2% HWCI from the eight qPR runs. Each column presents data from a different observer, and the four rows present estimated iconic memory decay functions obtained with 20, 50, 100, and 200 qPR trials. Agreement of the two methods and the precision of the estimated iconic memory decay functions with the corresponding number of qPR trials are summarized in Table 5.

Figure 14 shows that the iconic memory decay functions obtained with the qPR essentially overlapped with those obtained with the MCS. Agreement between the two methods was evaluated by the root mean squared error (RMSE) between the probability correct estimated from the two procedures at the eight cue delays in the MCS procedure:

$$RMSE = \sqrt{\frac{\sum_k \sum_j \sum_i (pc_{ijk}^{qPR} - pc_{ik}^{MSC})^2}{I \times J \times K}} \quad (14)$$

where pc_{ijk}^{qPR} is the probability correct at the i-th cue delay of the j-th run of the qPR procedure for the k-th observer, and pc_{ik}^{MCS} is the probability correct at the i-th cue delay for the k-th observer in the MCS procedure. The RMSE started at 0.149 after the first qPR trial and decreased to 0.063 after 100 qPR trials and to 0.049 after 200 qPR trials.

The precision of the qPR procedure is illustrated by the credible interval as a function of trial number in Figure 14. For each observer, the average 68.2% HWCI for each cue delay was obtained with 20, 50, 100, and 200 qPR trials. Results clearly show that the 68.2% HWCI of the estimated iconic memory decay functions decreased with trial number, reaching 0.055 probability correct after 100 qPR trials and 0.042 after 200 qPR trials.

Test–retest reliability was assessed with the Bland-Altman analysis (Bland & Altman, 1999) and the coefficient of repeatability (COR), which describes the 95% confidence limits for repeated measures. The CORs of the parameters of the iconic memory decay function, a_0, a_1, and τ, were 0.14, 0.05, and 0.28 (log10

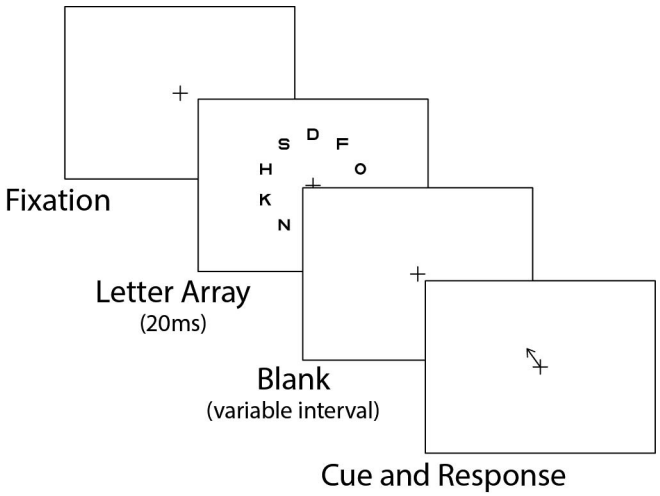

Figure 13. Illustration of the trial sequence in the psychophysical experiment. Each trial began with a fixation cross. The stimulus display contained 10 letters and was followed by a central cue. Observers were asked to report the cued letter. The cue remained on the screen until the observer responded.

units) after 200 qPR trials, respectively. The average COR of probability correct was 0.15, 0.14, and 0.12 for the three observers. The Bland-Altman plots in the upper row of Figure 15 show that the difference of parameter estimates between the first and second runs of the qPR procedure in each session fell within the 95% confidence limits (top and bottom dashed lines). The Bland-Altman plots on probability correct (the bottom row of Figure 15) also show that the estimated iconic memory decay functions from repeated runs of the qPR procedure were quite consistent.

Discussion

In the current study, we developed a quick partial report or qPR procedure based on a Bayesian adaptive framework to directly estimate the iconic memory decay function. In the qPR procedure, parameters of the exponential decay function are characterized by probability distributions. Starting with a prior distribution of the parameters, the qPR method selects and tests the observer at the most informative cue delay, by evaluating the stimulus space to find the cue delay that maximizes the expected information gain or minimizes the entropy of the parameters of the iconic memory decay function. It then updates the probability distribution of the parameters based on the observer's response by Bayesian inference (Kontsevich & Tyler, 1999; Lesmes et al., 2006, 2010; Watson & Pelli, 1983). The procedure is iterated until either the total number of trials reaches a set value or the precision of the parameter estimates reaches a certain criterion.

Compared to the conventional MCS procedure, the qPR uses a much broader stimulus range and a much finer stimulus sampling resolution. It estimates the whole shape of the iconic memory decay function with much less testing, since it concurrently measures an observer's performance across all different cue delays and utilizes all available information acquired during the experiment as well as prior knowledge about the functional form of the iconic memory decay function.

Results from the simulations and psychophysical experiment showed that the qPR method requires only 100 trials of data collection to measure the sensory memory decay function with reasonable accuracy and precision. Simulation studies suggest that only 100 trials are necessary to reach 0.026 average absolute bias and 0.070 precision (in probability correct). The method was also validated in a psychophysical experiment. Estimates of the sensory memory decay function obtained with 100 qPR trials showed good precision (68.2% HWCI = 0.055) and excellent agreement with those obtained with 1,600 trials of the MCS procedure (mean RMSE = 0.063). With the qPR procedure, reasonably precise estimates can be obtained in 5–10 min, which is significantly less than the typical one-hour test time for the conventional MCS method.

The prior in the qPR procedure can be informed by knowledge about the parameters of the iconic memory decay function as a function of the task, stimulus or test population (King-Smith et al., 1994). A proper informative prior can further speed up the estimation process. Researchers often use a broad bell-shaped distribution centered on the most probable value (e.g., Gaussian or hyperbolic secant distribution; King-Smith et al., 1994; Kontsevich & Tyler, 1999) to gain some benefit of prior knowledge but at the same time avoid some of the risks from using an inadequate prior. Alternatively, the prior could be a uniform distribution over the entire parameter space when researchers do not have much prior knowledge of the parameters of the iconic memory decay function. In our psychophysical experiment, we used a uniform prior distribution for each qPR parameter. A weakly informative prior based on prior knowledge of the parameters would make the measurement even more efficient if the test population is relatively homogeneous.

As life expectancy continues to increase, the number of people with AD rapidly grows each year. A recent study (Hebert, Weuve, Scherr, & Evans, 2013) shows that 11% and 32% among 65+, and 85+ year-olds, for a total of about 5 million people, have AD in the United States. The economic costs for AD, both direct and indirect, are enormous. Since the current medical treatments for AD could only relieve symptoms but not stop or reverse the progress of the disease, early detection and intervention is crucial (Sperling et al.,

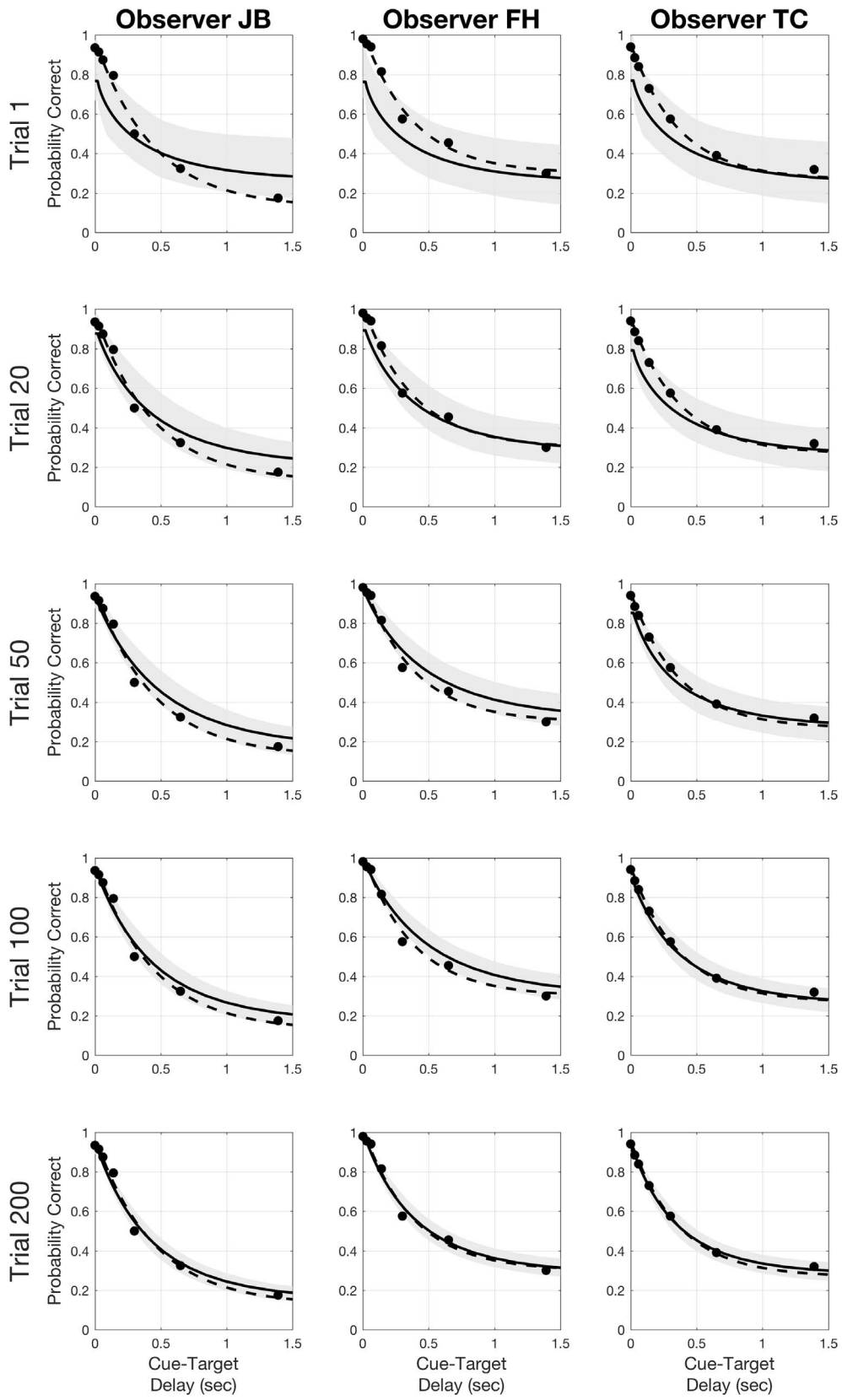

Figure 14. Results from the psychophysical experiment for three observers. Results from the two methods are presented in solid lines (qPR) and filled circles (MCS). Data obtained with MCS were fitted with the exponential decay function (Equation 1) using a maximum likelihood procedure. Best fitting functions are presented in dashed lines. The gray-shaded region shows the average 68.2% HWCI of eight qPR estimates. Different columns show results from different observers. Results after 20, 50, 100, and 200 qPR trials are shown in different rows.

	RMSE	68.2% HWCI
20 trials	0.110	0.095
50 trials	0.074	0.068
100 trials	0.063	0.051
200 trials	0.049	0.039

Table 5. Accuracy and precision of qPR in the psychophysical experiment.

2011). Without detection and interventions in the early stages of AD, the effects of treatment would be limited (Wimo & Winblad, 2006).

In current clinical practice, AD is diagnosed only after patients show obvious cognitive impairments accompanying difficulties in daily activities. However, precursors of AD might be presented long before it is clinically diagnosed, even before patients notice any observable behavioral impairments (Rowe et al., 2010). There have been major attempts to develop instruments for early AD detection, including neuroimaging techniques for biological markers (Fagan et al., 2006; Johnson, 2006; Mintun et al., 2006; Rowe et al., 2007) and neurocognitive tests (Rentz et al., 2013).

As shown by Lu et al. (2005), the fast decay of iconic memory might be a potential early sign of AD. The development of the qPR makes it possible to further evaluate the efficacy of fast decay of iconic memory as an early diagnosis signal of AD in clinical trials. In addition to the short test time, the qPR procedure is fully automatized. It assures more objective testing without requiring much effort from specialists in setting up the test and analyzing the results. Currently, most studies for early AD detection focus on neural signatures (e.g., deposition of peptide in brain). If the efficacy of fast decay of iconic memory as an early sign of AD is established, the qPR may be used to detect behavioral signals of AD without any special device: a generic personal computer is sufficient to run the qPR. It requires neither a special device, such as MRI or PET

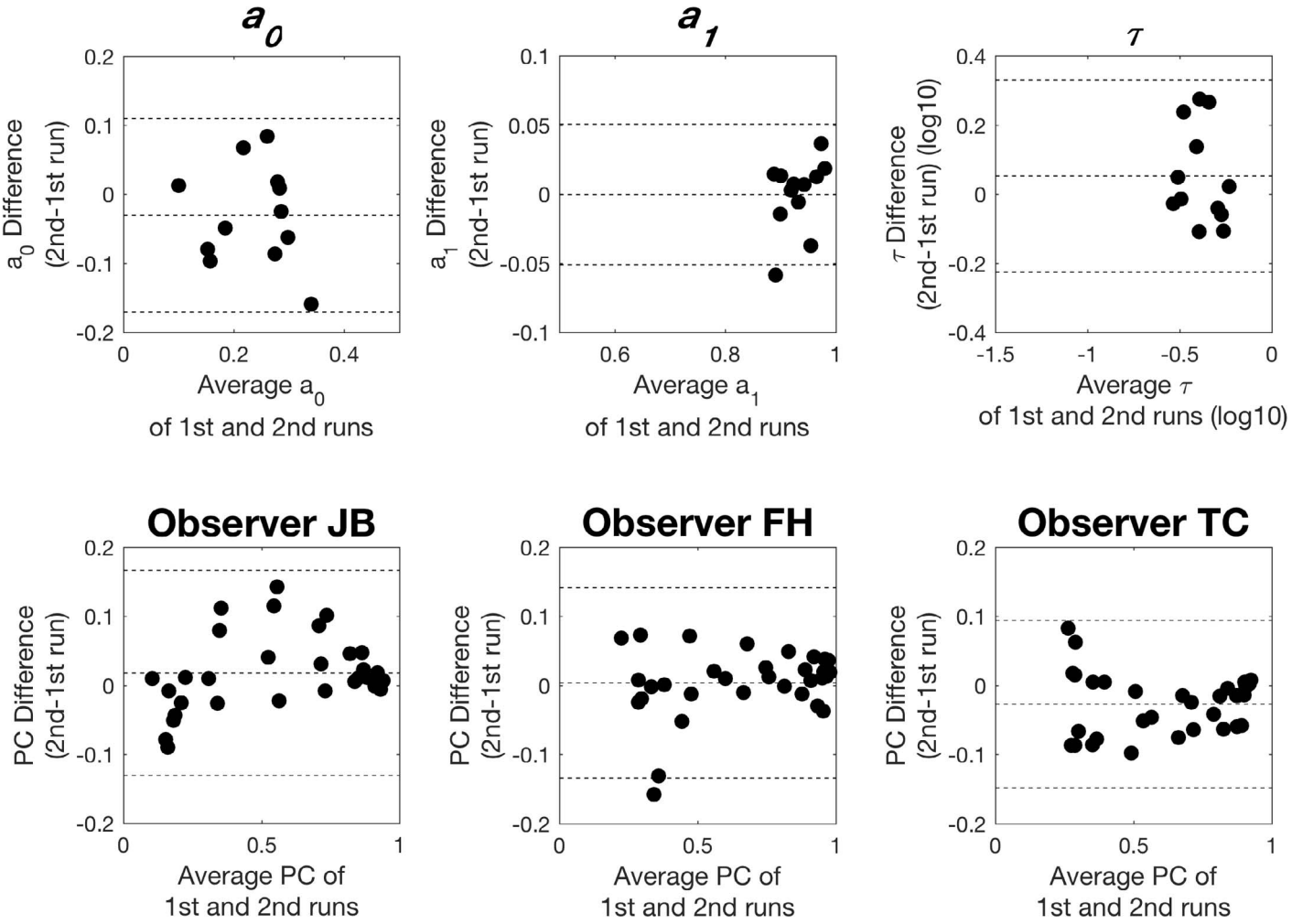

Figure 15. Test–retest repeatability: Bland-Altman plots on parameter estimates (upper row) and probability correct (bottom row). Dots represent the difference (second-run–first-run of the qPR in each session) plotted against the averages of the two measurements. Dashed lines represent the mean difference (middle), and at the limits of agreement (top and bottom).

machines, nor special equipment for biological analysis of blood or genome. Further development of the qPR method on mobile devices would make the test much easier and more convenient (Dorr, Lesmes, Lu, & Bex, 2013). The convenience and efficiency in terms of time, efforts, costs, and equipment could make qPR a highly valuable instrument for early AD detection.

In addition to be a potential tool in early AD detection, the qPR can also serve as a valuable tool in assessing iconic memory performance in cognitive tests. For example, the work by Miller et al. (2010) suggests that qPR may be used to assess psychometric intelligence.

The one-step-ahead search for the optimal test stimulus over multiple parameters requires intensive computations. Integration of modern computational algorithms, such as Markov Chain Monte Carlo (MCMC) sampling of the prior, into the qPR could reduce the computation time. MCMC methods are a class of algorithms that are based on sampling or estimating the posterior distributions as a function of the multidimensional parameter and multidimensional stimulus spaces. It is estimated that MCMC algorithms may reduce the computational load by a factor of 100 or more.

The qPR, described as an estimation procedure of iconic memory decay function, can also be implemented as a procedure to classify patients based on their performance and knowledge of different patient categories. In the case of discrimination, it is not the partial report parameters that are estimated, but the probability that the test observer is a member of a class defined by a prototypical partial report function that can signify normal or impaired iconic memory.

The qPR procedure, like most other adaptive procedures, is vulnerable to lapses at the beginning of each experiment. About 80–120 trials are necessary to recover from the impact of the early lapses trials. The performance of the method might be further improved through developments of more sophisticated procedures to detect and treat lapse trials. In addition, although one-step-ahead search—finding the "current best" stimulus for the next trial—has been proven to be efficient in many adaptive procedures, the greedy search algorithm is not necessarily the optimal strategy over the course of the whole experiment. Efficiency and robustness of methods could be improved by adopting global optimization or multiple-step-ahead strategy (Gu, Myung, Pitt, & Lu, 2013). The search can also incorporate considerations of other features of the test (e.g., total test time) in determining the optimal test sequence.

The qPR procedures herein assume a single functional form, but sometimes there could be two or more competitive models that could describe data well (e.g., exponential function vs. power function for the temporal property of memory decay). It would be useful to combine methods for selecting the best model among several different models with parameter estimation of each model into a single procedure (Cavagnaro, Myung, Pitt, & Kujala, 2010; Myung & Pitt, 2009). Hierarchical Bayesian modeling could provide even greater efficiency by better-informed priors across sessions or observers based on statistical dependency of data (Kim et al., 2014).

Keywords: adaptive procedure, iconic memory, partial report, Bayesian inference, information gain, Alzheimer's disease

Acknowledgments

This research was supported by the US National Institute of Health grants MH081018, EY017491, and EY021553, and Korean Ministry of Science, ICT and Future Planning grant IITP-2015-R0346-1008.

Commercial relationships: JB, LAL, and Z-LL have intellectual property interest in methods for measuring sensory memory. LAL and Z-LL have equity interest in a company (Adaptive Sensory Technology, Inc.) commercializing adaptive tests, including qCSF and qPR. LAL holds employment in Adaptive Sensory Technology.
Corresponding author: Zhong-Lin Lu.
Email: lu.535@osu.edu.
Address: Laboratory of Brain Processes (LOBES), Center for Cognitive and Brain Sciences & Center for Cognitive and Behavioral Brain Imaging, Departments of Psychology, The Ohio State University, Columbus, OH, USA.

Footnote

[1] This definition of average absolute bias is based on probability correct. It does not distinguish the differences associated with different baseline probability correct (e.g., the difference between 0.98 and 1.0 and the difference between 0.52 and 0.50).

References

Alcalá-Quintana, R., & García-Pérez, M. A. (2005). Stopping rules in Bayesian adaptive threshold estimation. *Spatial Vision*, *18*(3), 347–374. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16060232

Androutsopoulos, I., Paliouras, G., Karkaletsis, V., Sakkis, G., Spyropoulos, C. D., & Stamatopoulos, P. (2000). Learning to filter spam e-mail: A comparison of a naive Bayesian and a memory-based approach. In H. Zaragoza, P. Gallinari, & M. Rajman (Eds.), *Proceedings of the workshop "Machine learning and textual information access", 4th European conference on princples and practice of knowledge discovery in database* (pp. 1–13). Berlin, Germany: Springer Berlin Heidelberg.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation, Vol. 2.* (pp. 89–195). New York: Academic Press.

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*(1), 309–328. doi:10.1002/j.1538-7305.1961.tb03987.x.

Averbach, E., & Sperling, G. (1961). Short-term storage of information in vision. In C. Cherry (Ed.), *Symposium on information theory*. London, UK: Butterworth.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press, doi:10.1016/S0079-7421(08)60452-1.

Becker, M. W., Pashler, H., & Anstis, S. M. (2000). The role of iconic memory in change-detection tasks. *Perception*, *29*, 273–86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10889938

Berry, S. M., Carlin, B. P., Lee, J. J., & Muller, P. (2010). *Bayesian adaptive methods for clinical trials*. Boca Raton, FL: CRC Press.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*, 135–160.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. doi:10.1163/156856897X00357.

Cavagnaro, D. R., Myung, J. I., Pitt, M. A., & J. V. Kujala, (2010). Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural Computation*, *22*, 887–905. doi:10.1162/neco.2009.02-09-959.

Chong, M. S., & Sahadevan, S. (2005). Preclinical Alzheimer's disease: Diagnosis and prediction of progression. *Lancet Neurology*, *4*, 576–579. doi:10.1016/S1474-4422(05)70168-X.

Clayton, D., & Hills, M. (2013). *Statistical models in epidemiology*. Oxford, UK: Oxford University Press.

Cobo-Lewis, A. B. (1997). An adaptive psychophysical method for subject classification. *Perception & Psychophysics*, *59*, 989–1003. doi:10.3758/BF03205515.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, *27*, 183–228. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6992093

Dawid, A. P. (2005). Bayes' theorem and weighing evidence by juries. In R. Swinburne (Ed.), *Baye's theorem* (pp. 70–90). Oxford, UK: British Academy, doi:10.5871/bacad/9780197263419.001.0001.

Derman, C. (1957). Non-parametric up-and-down experimentation. *The Annals of Mathematical Statistics, 28*(3), 795–798, doi:10.1214/aoms/1177706895.

Dick, A. O. (1974). Iconic memory and its relation to perceptual processing and other memory mechanisms. *Perception & Psychophysics*, *16*, 575–596. doi:10.3758/BF03198590.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology. General*, *109*, 75–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6445405

Dixon, R. A., Garrett, D. D., Lentz, T. L., MacDonald, S. W., Strauss, E., & Hultsch, D. F. (2007). Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. *Neuropsychology*, *21*, 381–99. doi:10.1037/0894-4105.21.3.381.

Dorr, M., Lesmes, L. A., Lu, Z.-L., & Bex, P. J. (2013). Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Investigative Ophthalmology & Visual Science*, *54*, 7266–7273. [PubMed] [Article]

Dorr, M., Wille, M., Viulet, T., Sanchez, E., Bex, P. J., Lu, Z. L., & Lesmes, L. (2015). Next-generation vision testing: The quick CSF. *Current Directions in Biomedical Engineering*, *1*(1), 131–134.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, *70*, 193–242. doi:10.1037/h0044139.

Fagan, A. M., & Mintun, M. A., Mach, R. H., Lee, S.-Y., Dence, C. S., Shah, A. R., ... Holtzman, D. M. (2006). Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Abeta42 in humans. *Annals of Neurology*, *59*, 512–519. doi:10.1002/ana.20730.

Gu, H. (2012). *Graphic-processing-units based adaptive*

parameter estimation of a visual psychophysical model. (Master's thesis). Retrieved from https://etd.ohiolink.edu/

Gu, H., Kim, W., Hou, F., Lesmes, L. A., Pitt, M. A., Lu, Z., & Myung, J. I. (2016). A hierarchical Bayesian approach to adaptive vision testing: A case study with the contrast sensitivity function. *Journal of Vision*, *16*(6):15, 1–17, doi:10.1167/16.6.15. [PubMed] [Article]

Gu, H., Myung, J. I., Pitt, M. A., & Lu, Z. (2013). Bayesian adaptive estimation of psychometric slope and threshold with differential evolution not all designs are equally informative. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 2452–2457). Austin, TX: Cognitive Science Society.

Hebert, L. E., Weuve, J., & Scherr, P. A., & Evans, D. A., (2013). Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. *Neurology*, *80*, 1778–83. doi:10.1212/WNL.0b013e31828726f5.

Hou, F., Huang, C.-B., Lesmes, L., Feng, L.-X., Tao, L., Zhou, Y.-F., & Lu, Z.-L. (2010). qCSF in clinical application: Efficient characterization and classification of contrast sensitivity functions in amblyopia. *Investigative Ophthalmology & Visual Science*, *51*, 5365–5377. [PubMed] [Article]

Hou, F., Lesmes, L., Bex, P., Dorr, M., & Lu, Z.-L. (2015). Using 10AFC to further improve the efficiency of the quick CSF method. *Journal of Vision, 15*(9)2, 1–18, doi:10.1167/15.9.2. [PubMed] [Article]

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.

Johnson, K. A. (2006). Amyloid imaging of Alzheimer's disease using Pittsburgh Compound B. *Current Neurology and Neuroscience Reports*, *6*, 496–503. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/170742 85

Jonides, J., Irwin, D., & Yantis, S. (1982). Integrating visual information from successive fixations. *Science*, *215*, 192–194. doi:10.1126/science.7053571.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics, 49*(3), 227–229. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2011460.

Keele, S. W., & Chase, W. G. (1967). Short-term visual storage. *Perception & Psychophysics*, *2*, 383–386. doi:10.3758/BF03210076.

Kesten, H. (1958). Accelerated stochastic approximation. *The Annals of Mathematical Statistics, 29*(1), 41–59, doi:10.1214/aoms/1177706705.

Kim, W., Pitt, M. A., Lu, Z., Steyvers, M., & Myung, J. I. (2014). A hierarchical adaptive approach to optimal experimental design. *Neural Computation*, *26*, 2465–2492.

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, *34*, 885–912. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8160402

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*, 2729–2737. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10492833

Kruschke, J. (2011). *Doing Bayesian data analysis: A tutorial with R and BUGS*. Burlington, MA: Academic Press.

Kujala, J., & Lukka, T. (2006). Bayesian adaptive estimation: The next dimension. *Journal of Mathematical Psychology*, *50*, 369–389. doi:10.1016/j.jmp.2005.12.005.

Kujala, J. V., Richardson, U., & Lyytinen, H. (2010). A Bayesian-optimal principle for learner-friendly adaptation in learning games. *Journal of Mathematical Psychology*, *54*, 247–255. doi:10.1016/j.jmp.2009.10.001.

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*, 1279–1292. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11800457

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*(2), 467–477. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5541744.

Lesmes, L. A., Jeon, S.-T., Lu, Z.-L., & Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method. *Vision Research*, *46*, 3160–76. doi:10.1016/j.visres.2006.04.022.

Lesmes, L. A., Lu, Z.-L., Baek, J., & Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: The quick CSF method. *Journal of Vision*, *10*(3):17, 1–21, doi:10.1167/10.3.17. [PubMed] [Article]

Lesmes, L. A., Lu, Z.-L., Baek, J., Tran, N., Dosher, B. A., & Albright, T. D. (2015). Developing Bayesian adaptive methods for estimating sensitivity thresholds (d′) in Yes-No and forced-choice tasks. *Frontiers in Psychology, 6,* 1070, doi:10.3389/fpsyg.2015.01070.

Long, G. M. (1980). Iconic memory: A review and

critique of the study of short-term visual storage. *Psychological Bulletin*, *88*, 785–820. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7003642

Lu, Z.-L., & Dosher, B. (2013). *Visual psychophysics: From laboratory to theory*. Cambridge, MA: The MIT Press.

Lu, Z.-L., Neuse, J., Madigan, S., & Dosher, B. A. (2005). Fast decay of iconic memory in observers with mild cognitive impairments. *Proceedings of the National Academy of Sciences, USA*, *102*, 1797–1802. doi:10.1073/pnas.0408402102.

Miller, R., Rammsayer, T. H., Schweizer, K., & Troche, S. J. (2010). Decay of iconic memory traces is related to psychometric intelligence: A fixed-links modeling approach. *Learning and Individual Differences*, *20*, 699–704.

Mintun, M. A., Larossa, G. N., Sheline, Y. I., Dence, C. S., Lee, S. Y., Mach, R. H., ... Morris, J. C. (2006). [11C]PIB in a nondemented population: Potential antecedent marker of Alzheimer disease. *Neurology*, *67*, 446–452. doi:10.1212/01.wnl.0000228230.26044.a4.

Myung, J. I., & Pitt, M. A. (2009). Optimal experimental design for model discrimination. *Psychological Review*, *116*, 499–518. doi:10.1037/a0016104.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. doi:10.1163/156856897X00366.

Peripheral and Central Nervous System Drugs Advisory Committee. (2001). *Vascular dementia*. Gaithersburg, MD: U.S. Food and Drug Administration, http://www.fda.gov/ohrms.dockets/ac/01/briefing/3724b2_01_VasDementia.pdf.

Prins, N. (2013). The psi-marginal adaptive method: How to give nuisance parameters the attention they deserve (no more, no less). *Journal of Vision*, *13*(7):3, 1–17, doi:10.1167/13.7.3. [PubMed] [Article]

Rentz, D. M., Parra Rodriguez, M. A., Amariglio, R., Stern, Y., Sperling, R., & Ferris, S. (2013). Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: A selective review. *Alzheimer's Research & Therapy*, *5*(6), 1–10, doi:10.1186/alzrt222.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The Annals of Mathematical Statistics, 22*(3), 400–407, doi:10.1214/aoms/1177729586.

Rothman, K. J., & Greenland, S. (1998). *Modern Epidemiology* (2nd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Rowe, C. C., Ellis, K. A., Rimajova, M., Bourgeat, P., Pike, K. E., Jones, G., ... Villemagne, V. L. (2010). Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiology of Aging*, *31*, 1275–1283. doi:10.1016/j.neurobiolaging.2010.04.007.

Rowe, C. C., Ng, S., Ackermann, U., Gong, S. J., Pike, K., Savage, G., ... Villemagne, V. L. (2007). Imaging beta-amyloid burden in aging and dementia. *Neurology*, *68*, 1718–1725. doi:10.1212/01.wnl.0000261919.22630.ea

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423. doi:10.1145/584091.584093.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5,* 19–31, doi:10.1177/001872086300500103.

Sperling, G. (1967). Successive approximations to a model for short-term memory. *Acta Psychologica*, *27*, 285–292. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6062221

Sperling, R., Aisen, P. S., Beckett, L., Bennett, D., Craft, S., Fagan, A. M., ... Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*, 280–292. doi:10.1016/j.jalz.2011.03.003.

Tanner, T. (2008). Generalized adaptive procedure for psychometric measurement. *Perception, 37*(1 Supplement), 93, doi:10.1177/03010066080370S101.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America, 41*(4A), 782, doi:10.1121/1.1910407.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, *35*, 2503–2522. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8594817

Tyrrell, R. A., & Owens, D. A. (1988). A rapid technique to assess the resting states of the eyes and other threshold phenomena: The modified binary search (MOBS). *Behavior Research Methods, Instruments, & Computers, 20*(2), 137–141, doi:10.3758/BF03203817.

Watson, A B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113–120. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6844102

Wimo, A., & Winblad, B. (2006). Health economics of severe dementia. In A. Burns & B. Winblad (Eds.), *Severe dementia* (pp. 227–236). Chichester, UK: John Wiley & Sons, Ltd., doi:10.1002/0470010568.ch19.

Yang, Z., & Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. *Molecular Biology and Evolution*, *14*, 717–724. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/921474