

Detecting chance: A solution to the null sensitivity problem in subliminal priming

JEFFREY N. ROUDER, RICHARD D. MOREY, PAUL L. SPECKMAN, AND MICHAEL S. PRATTE
University of Missouri, Columbia, Missouri

In many paradigms, the persuasiveness of subliminal priming relies on establishing that stimuli are undetectable. The standard significance test approach is ill-suited as null results may reflect either truly undetectable stimuli or a lack of power to resolve weakly detectable stimuli. We present a novel statistical model as an alternative. The model provides for estimates of the probability that each individual is truly at chance. Researchers may select individuals for whom there are sufficiently high probabilities of true undetectability. The model is hierarchical, and estimation is done within the Bayesian framework.

Although there are numerous claims of subliminal priming, the topic remains controversial. We propose a new model for assessing the subliminality of stimuli. We illustrate the model in the number priming paradigm of Dehaene et al. (1998). In this paradigm, participants classify target digits (e.g., “8”) as either greater-than or less-than five. Preceding the target digit is a briefly presented and subsequently masked digit which serves as a prime. Prime–target pairs are considered congruent if both have the same relation to five (e.g., 2 and 3 are congruent because both are less than 5) and incongruent otherwise. The claim for subliminality has been made when response times are faster to congruent prime–target pairs than to incongruent prime–target pairs, yet primes are not detectable (e.g., Dehaene et al., 1998; Van Opstal, Reynvoet, & Verguts, 2005).

One of the most vexing problems in this line of research is establishing that primes are truly not detectable. Typically, researchers run a set of trials in which they simply ask participants to detect the presence of the prime (or to discriminate among primes). Performance is quite low and often statistically indistinguishable from chance. Although the statistical analysis is computationally straightforward, the logic of claiming that the primes are undetectable is problematic. Reingold and Merikle (1988) popularized the following argument: The appropriate null hypothesis in this application is that detection performance is at chance. Evidence for undetectability is a failure to reject this null. Unfortunately, failure to reject this null may come about from insufficient power to resolve low but above-chance performance. Given that most subliminal priming effects are relatively small (see Draine & Greenwald, 1998), it is plausible that they result from superliminal priming with weakly detectable stimuli.

Current claims about the subliminality of a prime roughly fall into one of two classes. In the first class, re-

searchers test whether a specific participant is able to perceive a prime. In the second class, researchers test whether a group of participants, en masse, can perceive a prime. We address these claims in turn.

The simplest means of testing for subliminality for a specific participant is to test whether the 95% confidence interval around the observed performance score includes the chance baseline. Consider the case in which performance is measured as accuracy in a two-choice task with a chance baseline of .5. Figure 1, top left panel, shows the probability of concluding subliminality for various above-chance levels of performance. The results are discouraging. For example, if the sample size is 50 trials (line A) and the participant has a true accuracy of .6, the probability of concluding the prime is subliminal is .65 (marked with a “+”). Even at larger samples, this error rate remains high. For example, at 400 trials (third line from the left), true accuracy scores of .55 will be misdiagnosed as subliminal with a rate over 0.45 (marked with a “O”).

Alternatively, researchers have claimed subliminality by decreasing duration, in fine increments, until a threshold criterion was reached. For example, Dagenbach, Carr, and Wilhelmsen (1989) decreased duration in increments of 5 msec until participants responded correctly on no more than 23 out of 40 trials. Figure 1, top center panel, shows why this approach is problematic. For Dagenbach et al.’s criterion, there is still a high probability of concluding subliminality when true performance is substantially above chance. Another alternative is provided by Greenwald, Klinger, and Liu (1989), who selected participants whose observed performance was below chance (i.e., accuracy was less than .5). The approach, however, is still susceptible to the same critique. Figure 1, top right panel, shows the probability of accepting subliminality as a function of sample size. Once again, if true values are near .5, the method is susceptible to mistakenly concluding subliminality.

J. N. Rouder, rouderj@missouri.edu

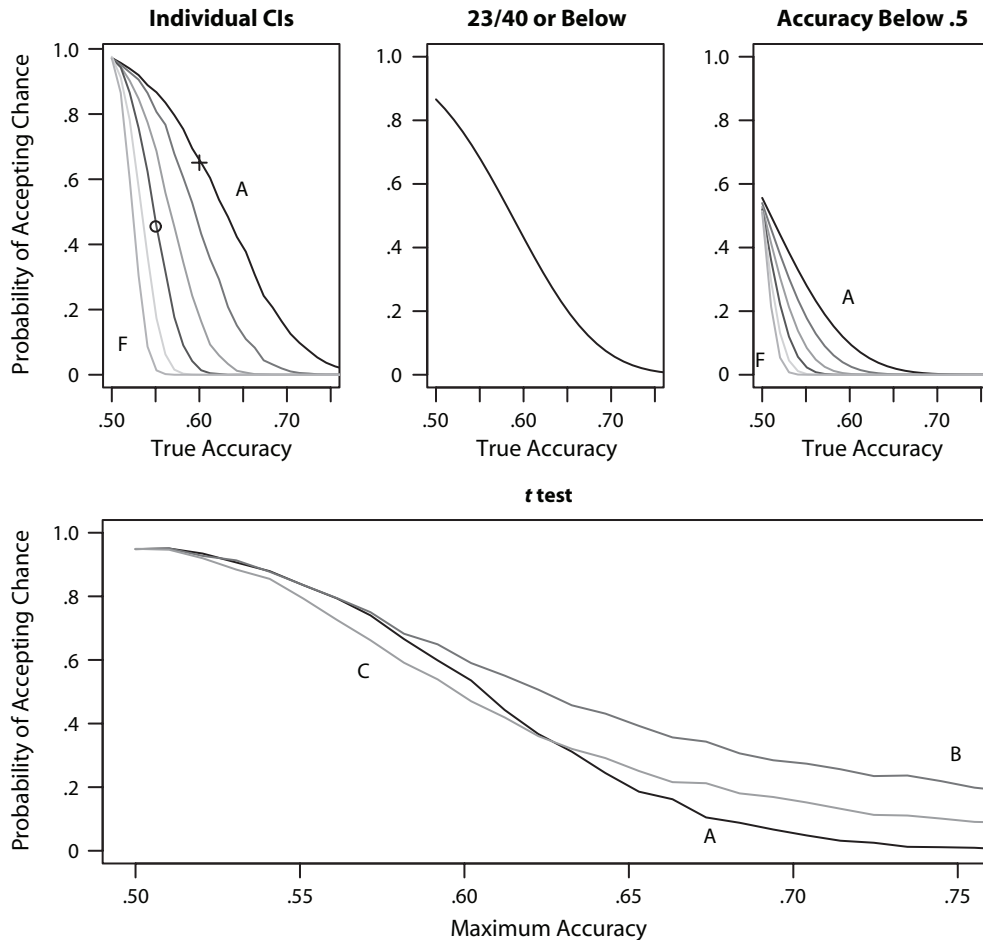


Figure 1. Probability of accepting chance performance as a function of true accuracy. Top left: 95% confidence intervals method. Lines A–F denote sample sizes 50, 100, 200, 400, 800, and 1,600, respectively. Top middle: Dagenbach, Carr, and Wilhelmsen's (1989) criterion of 23 of 40 correct responses. Top right: Greenwald, Klinger, and Liu's (1989) selection of below-chance performance. Lines A–F denote sample sizes 50, 100, 200, 400, 800, and 1,600, respectively. Bottom: *t* test method. Lines A–C represent the sample sizes from Murphy and Zajonc's (1993) Experiment 1 and from Dehaene et al.'s (1998) Task 1 and Task 2, respectively.

The second class of claims about subliminality is made for a group of participants. Typically, the collection of individuals' accuracy scores are submitted to a *t* test. The null hypothesis is that the true group average is at chance and a failure to reject this null is taken as evidence for subliminality of the prime for all participants. We highlight two influential articles: Dehaene et al.'s (1998) investigation of the neural correlates of subliminal priming, and Murphy and Zajonc's (1993) demonstration of affective subliminal priming from emotive faces. Dehaene et al. asked 6 participants to discriminate number primes from blank fields (96 trials) and 7 participants to discriminate number primes from letter strings (112 trials). For prime durations of 33 msec or less, they reported that mean accuracies in both tasks were not significantly different from chance. Murphy and Zajonc (1993) asked 32 participants to identify whether a prime face was happy or scowling on 12 trials. They too found that mean accuracy did not differ significantly from chance. We assessed the probability of mistakenly accepting subliminality with a *t* test through simulations. Each participant

was assumed to have some degree of superliminal performance. True accuracies were uniformly distributed between .5 and an upper bound. Figure 1, bottom, shows the probability of accepting the null as a function of the upper bound for the sample sizes in these experiments. Even for high values of this upper bound, there is a substantial probability of mistakenly concluding subliminality.

Figure 1 presents a troubling picture; it shows the difficulty of determining subliminality. Conventional null hypothesis tests are ill-suited here because they do not offer appropriate safeguards from mistakenly concluding subliminal performance. In this article, we provide a new model which provides these safeguards. Although the model has some shortcomings, it is a dramatic improvement over current techniques.

MODEL DEVELOPMENT

To explicate the model, we consider accuracy in a two-alternative forced-choice task with a chance baseline of

.5. Extensions to other tasks and measures are possible but not covered. It is reasonable to assume that each individual's performance reflects a true probability of a correct response (denoted by p_i). The number of correct responses for the i th participant, denoted y_i , is modeled conventionally as a binomial:

$$y_i \sim \text{binomial}(p_i, N_i), \tag{1}$$

where N_i is the number of trials for i th participant. Some individuals may be truly at chance ($p_i = .5$), whereas others may be able to detect the prime ($p_i > .5$).

Each participant is assumed to have a true latent ability, denoted by x_i . True probabilities (p_i) are related to latent abilities:

$$p_i = \begin{cases} .5, & x_i \leq 0, \\ \Phi(x_i), & x_i > 0, \end{cases} \tag{2}$$

where Φ is the standard normal cumulative distribution function that relates z scores to probabilities. Figure 2 shows an example of the model. The lower axis shows individuals' true latent abilities. The upper axis shows corresponding values of true probability p_i . Negative values of x_i correspond to subliminality; positive values correspond to superliminality.

Three hypothetical participants are shown as examples in Figure 2. They have latent abilities $x_1 = -.4, x_2 = -.2$, and $x_3 = .3$, respectively. Whereas the first 2 participants have negative latent abilities, their true probabilities are $p_1 = p_2 = .5$. The true probability for the 3rd participant is the probit transform of x_3 ; that is, $p_3 = \Phi(x_3) = \Phi(.3) = .62$. Positive values of x_1 are interpreted as probit scores of performance. Negative values are interpreted as the amount of increase in visibility needed for above-chance performance. Latent abilities depend on stimulus factors. For example, the same participant will show higher latent

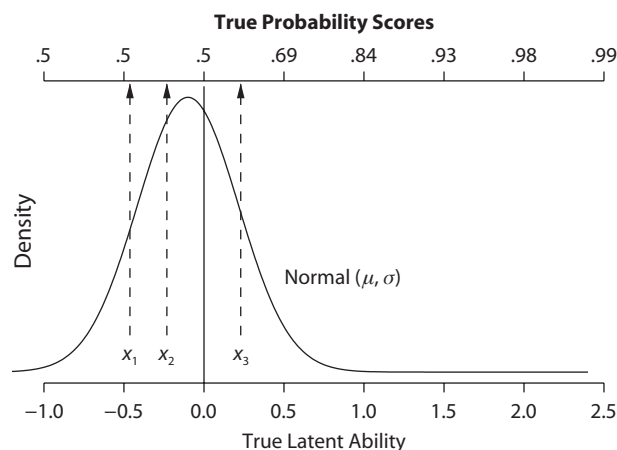


Figure 2. The mass at chance model. The top and bottom axes depict participants' latent ability (x_i) and true probability (p_i), respectively. The curve shows the population distribution of latent abilities. Sixty-two percent of the area under this curve is to the left of zero, indicating that 62% of the population is at chance. The three vertical lines denote latent abilities and corresponding probabilities for three hypothetical participants.

ability to primes presented for 33 msec than to those presented for 22 msec.

One novel feature of the model is that some proportion of participants have true probabilities at chance. For these participants, the prime is undetectable on every trial. The model, therefore, makes a strong psychological commitment: There are stimuli that are so weak that they are imperceptible on all trials. To highlight this commitment, we name the model the *mass at chance model* (MAC). This commitment is not present in classic psychometric theory (e.g., 3PL model; Nunnally & Bernstein, 1994). In these models, true probability at chance is an unreachable lower bound.

Theorists of near-liminal processing have debated the existence of thresholds. The debate is usually couched in terms of threshold models (Blackwell, 1953; Egan, 1975; Luce, 1963) or signal-detection models (Green & Swets, 1966). In threshold models, perceptual information on each trial is dichotomous. In signal detection, perceptual information is, alternatively, graded. The MAC model assumes that each participant has a threshold. This threshold, however, differs from the threshold in information-processing models in that it does not describe perception on a trial-by-trial level. Instead, it describes performance across all trials. MAC is certainly compatible with traditional threshold models; detection on any trial is possible only when $x_i > 0$. MAC is also compatible with signal detection models: d' is non-zero only when $x_i > 0$. Various trial-by-trial information-processing models may be nested within the MAC model with sensitivity parameters subject to the population constraints imposed by the MAC hierarchy.

The MAC model is conveniently analyzed in the Bayesian framework. In this framework, the analyst computes *posterior distributions*, or distributions of parameters reflecting the influence of data (Lee & Wagenmakers, 2005). The main quantities of interest are the posterior distributions of participants' latent abilities ($x_i | \text{data}$). Whereas subliminality corresponds to negative values of x_i , we compute the posterior probability $Pr(x_i < 0 | \text{data})$. We consider the i th participant at chance if this posterior probability is above a criterion. The choice of this criterion may be set by convention to .95. The analyst's decision rule is, therefore,

$$Pr(x_i < 0 | \text{data}) \geq .95 \rightarrow \text{Conclude the participant is at chance.} \tag{3}$$

$$Pr(x_i < 0 | \text{data}) < .95 \rightarrow \text{Conclude the participant is above chance.} \tag{4}$$

The MAC model mitigates the null sensitivity problem. In the MAC model, small sample sizes result in highly variable posterior estimates of x_i . Consequently, there tends to be more mass above $x_i > 0$. Lowering sample sizes makes it more difficult to claim subliminality, providing a safeguard against underpowered designs.

In Bayesian analysis, priors must be placed on parameters—in this case, on latent abilities x_i . The choice of prior on x_i induces a corresponding prior on p_i , and these correspondences are shown in Figure 3. The left panel shows the prior on p_i when the prior on x_i is flat, $x_i \sim \text{normal}(\mu = 0$,

$\sigma^2 = \infty$). Half of the mass is placed on $p_i = .5$; half is placed on $p_i = 1.0$; and an infinitesimal amount is placed on all values between .5 and 1.0. The center and right panel shows the corresponding priors for other choices. The center panel shows the prior on p_i when a standard normal prior is placed on x_i ; that is, $x_i \sim \text{normal}(\mu = 0, \sigma^2 = 1)$. The right panel shows the prior on p_i when the prior on x_i is normal ($\mu = 0, \sigma^2 = .05$). This prior reflects the reasonable belief that if half of the people cannot truly see the stimulus, then many of the remaining are going to cluster near chance.

Ideally, the choice of prior should not unduly affect the results in analysis with reasonable sample sizes. Unfortunately, in this application, it does. The posterior quantity $Pr(x_i < 0 \mid \text{data})$ can vary greatly for these three priors. Consider a participant who responds correctly on 140 of 288 trials. The posterior probability that this participant is at chance is $\approx 1, .95$, and $.64$, for values of $\sigma^2 = \infty, 1$, and $.05$, respectively.

The dependence on the posterior on the three priors above motivates the need for an additional level of complexity. We use a *hierarchical prior* on x_i :

$$x_i \sim \text{normal}(\mu, \sigma^2). \tag{5}$$

In this prior, parameters μ and σ^2 are estimated rather than assumed. These parameters are population-level parameters that describe how latent ability is distributed. The curve in Figure 2 denotes a hypothetical distribution of latent abilities. This prior is termed *hierarchical* because individuals' latent abilities are nested within a population-level distribution. Priors are needed for μ and σ^2 , but the specification of these is less critical and easier to justify than the previous nonhierarchical priors. Specification of the priors and an overview of analysis is found in the Appendix. Detailed development is found in a supplement at www.missouri.edu/~pcl/papers. Source code for the analysis, in both the C and R languages, and an easy-to-use Windows executable may be found at www.missouri.edu/~pcl/code.

Equations 1–5 specify a nonlinear, hierarchical Bayesian model. In previous applications, we have advocated

this class of models for paradigms in which researchers believe that a nonlinear model is appropriate. The advantage of this class of models is that it provides for accurate and powerful inference and does not require researchers to aggregate individuals for group-level tests. We have provided a number of nonlinear hierarchical models for psychological applications (Lu, Speckman, Sun, Rouder, & Morey, 2007; Rouder, Lu, Morey, Sun, & Speckman, in press; Rouder, Lu, Speckman, Sun, & Jiang, 2005; Rouder, Lu, Sun, et al., in press; Rouder, Sun, Speckman, Lu, & Dzhou, 2003) as well as a tutorial for psychologists (Rouder & Lu, 2005).

EXPERIMENT

We used the MAC model to assess the existence of subliminal priming. Our design is a near replication of Van Opstal et al. (2005); the structure of a trial is shown in Figure 4. For the *prime-identification task*, participants indicated the less-than-five status for primes. For the *target-identification task*, participants did the same for targets.

Method

Participants. Twenty-seven University of Missouri, Columbia undergraduates served as participants.

Design. Prime–target congruence (congruent/incongruent) served as the main independent variable. Prime identification performance and the priming effect in target identification served as the main dependent variables.

Procedure. The structure of each trial is shown in Figure 4. The primes were the digits 2, 3, 4, 6, 7, and 8; targets were the digits 1, 4, 6, and 9. Primes, targets, and masks were presented in white against a black background. The black background had a luminance of $.5 \text{ cd/m}^2$. Displays with digits had luminances that varied between 2.8 cd/m^2 and 4.5 cd/m^2 , depending on the number of pixels illuminated in the digit. Masks, which consisted of displays with 5 “#” characters, had a luminance of 16.5 cd/m^2 . Digits subtended approximately $.2^\circ$ of arc horizontally and twice that vertically. All pairings of primes and targets occurred equally often in each task for each participant. Each participant first observed 576 trials in the target-identification task followed by 288 trials in the prime identification task.

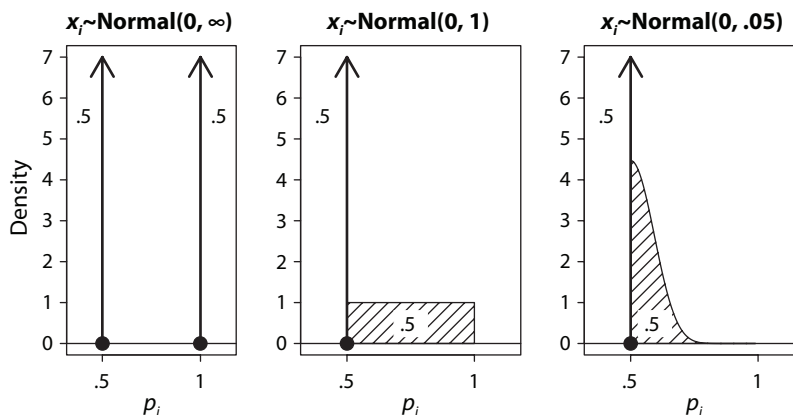


Figure 3. Relationship between priors on x_i and those on p_i . The lines extending upward indicate that half of the mass is concentrated at a point.

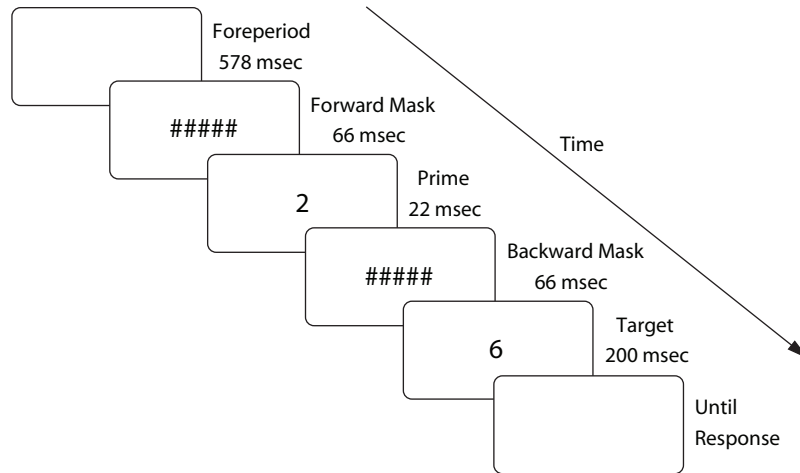


Figure 4. The structure of a trial.

Results

The resulting accuracies for the prime identification task are shown in the top panel of Figure 5. The lower panel shows MAC model estimates of the posterior probabilities that $x_i < 0$. Three of the 27 participants have posterior probabilities greater than the .95 criterion, and are considered as performing at chance (open circles). An additional two have posterior probabilities between .946 and .95, and we provisionally consider them as well. The remaining 22 participants are shown with filled circles and are inappropriate for evaluating subliminal priming. Although many of these participants may be truly at

chance, we do not have sufficient evidence from the data to conclude this.

One of the conventional analyses of these data is to test whether the population mean of proportion correct scores is different from the chance baseline of .5. A *t* test on the scores y_i/N_i yields a rejection of the null hypothesis [$t(26) = 2.724, p < .05$], although this result is leveraged by the two high-performing outliers. The advantage of the MAC model for this sample is that it allows for the selection of some of the 27 participants as not being able to detect the prime. These participants can then be used to test for the existence of subliminal priming.

The priming effect is shown in Table 1. There is a significant 8 msec effect across all 27 participants [$t(26) = 3.64, p < .05$].¹ Although the priming effect holds numerically for people selected as at chance (whether the criterion was .95 or .946), it is not significant. The MAC model reveals in this case that more participants are needed in order to assess the subliminal priming effect. Our experiment, however, is larger than most in the number priming literature.

EFFECTS OF REDUCING SAMPLE SIZE

One element of conventional significance tests that we see as especially problematic is the effect of reducing sample size. In these tests, reducing sample size increases the probability of claiming subliminality (from lack of power). This problem is not characteristic of the MAC model. The following simulation demonstrates that increasing the number of observations per participant leads to more certainty in analysis. We simulated data from 40 participants with half at chance and with half somewhat above chance ($\mu = 0, \sigma^2 = .18$). The resulting MAC posterior probabilities are shown in Figure 6. In one simulation, there were 100 observations per hypothetical participant (left panel); in the other, there were 1,000 observations per hypothetical participant (right panel). Of the 20 hypothetical participants truly at chance, 0 and 5 participants were identified at chance in the simulation with 100 and 1,000 samples, respectively. The MAC model is conservative,

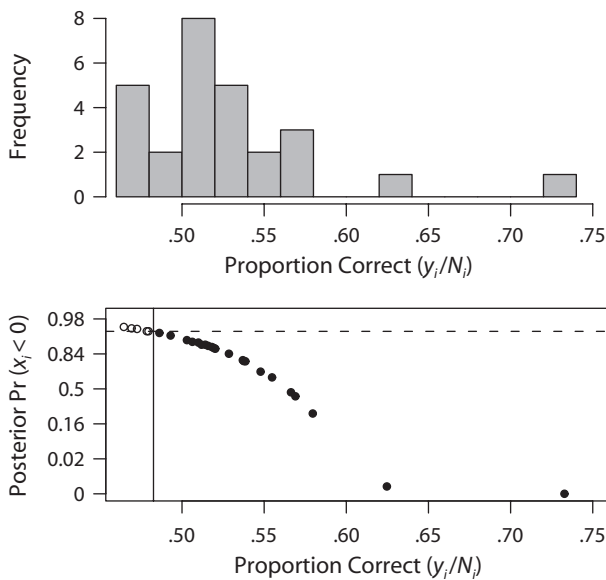


Figure 5. Top: Histogram of accuracy of the 27 participants. Bottom: MAC model estimates of the posterior probability that $x_i < 0$ as a function of accuracy. Open circles denote probabilities greater than .945, indicating some evidence to consider the participant at chance. Filled circles denote probabilities below .945, indicating insufficient evidence to consider the participant at chance.

Table 1
Effect of Prime Congruence

MAC Model Criterion	No. of Participants	Mean Response Time (msec)			<i>p</i> Value From <i>t</i> Test
		Congruent	Incongruent	Difference	
None	27	508	516	8	.001
.945	5	526	539	13	.10
.95	3	496	512	16	.26

but it is conservative in a direction that prevents concluding a participant is at chance for lack of evidence.

CAVEATS FOR APPLICATION

Although we have emphasized the advantages of the MAC model, there is a problem in application that demands attention. It is reasonable to expect that if the rules of Equations 3 and 4 are used, then, in the long run, a minimum of 95% of the selected participants will truly be at chance. Unfortunately, this expectation holds only asymptotically—that is, in the limit of either a large number of participants or a large number of trials per participant.

To assess the accuracy of decisions in reasonably sized experiments, we ran several simulation studies. Simulations began with the specification of true values for μ and σ^2 . For each replicate, we first sampled true values of x_i (in accordance with Equation 5) and then sampled observed accuracy scores (in accordance with Equation 1). For each replicate, we analyzed the data with the MAC model. Simulations were performed with 288 trials per participant, which represents a reasonable lower bound. The number of participants was either 10 per replicate or 30 per replicate. Table 2 shows the proportion of participants classified as at chance who were misclassified. Values greater than .05 are troubling. They occur for experiments with small numbers of participants and small true values of σ^2 . Therefore,

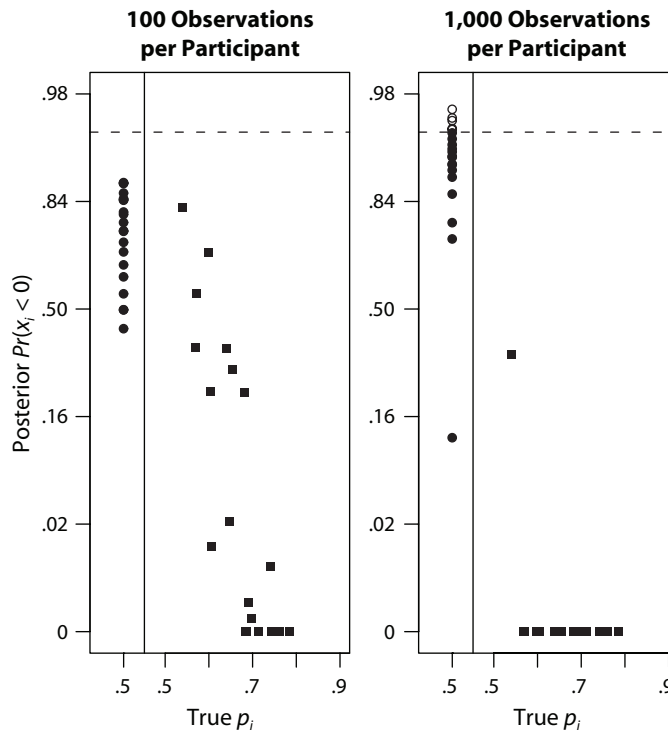


Figure 6. Effect of sample size. Analysis of simulated data in which half of 40 hypothetical participants are truly at chance. Plots show estimated posterior probabilities of chance performance as a function of true probabilities. Left and right plots are for sample sizes of 100 and 1,000 observations per participant, respectively. Open and filled symbols denote probabilities above and below the .95 criterion, respectively. Circles denote estimates for which true probabilities are at chance; squares denote estimates for which true probabilities are above chance. Increasing the sample size provides for more extreme posterior probabilities, which, in turn, provides for more precise classification of participants as at or above chance.

Table 2
MAC Model Simulations

Participants	True μ	True σ^2	Errors
30	-.2	.04	.09
30	0	.04	.10
30	.1	.04	.08
30	-.6	.36	.03
30	0	.36	.04
30	.3	.36	.05
10	-.2	.04	.10
10	0	.04	.21
10	.1	.04	.25
10	.2	.04	.08
10	-.6	.36	.05
10	0	.36	.09
10	.3	.36	.07

we recommend the following two guidelines: (1) Samples should consist of at least 30 participants observing about 300 trials. (2) Researchers should consider bolstering their claims with simulations. The simulation with 30 participants, $\mu = -.6$, and $\sigma^2 = .36$, which has low error, roughly corresponds to values obtained in the experiment. We have provided an easy-to-use Windows executable that performs these simulations (www.missouri.edu/~pcl/code).

The accuracy of MAC decisions is improved when σ^2 is accurately estimated. Unfortunately, it is difficult to estimate this parameter when many people have latent abilities (x_i) below zero. To provide for greater accuracy, we are currently developing MAC models for multiple stimulus durations. The longer durations will lead to above-chance, though not ceiling, levels of performance. These durations may be used for accurate estimation of σ^2 , which may aid in the assessment of subliminality at shorter durations.

A second caveat for application concerns the possibility that the prime identification task underestimates latent prime identifiability. Target identification is an easy task, and participants retain motivation and attention. Prime identification, however, is much more difficult. Pratte (2007) found that participants' motivation was substantially lower in prime than in target identification. Furthermore, when easy-to-identify primes were intermixed with near-liminal ones, identification of near-liminal primes increased markedly. Pratte concluded that the easy-to-identify primes raise motivation in the task. The MAC model does not address these concerns. Users still bear the burden of equating characteristics across the tasks (cf. Reingold & Merikle, 1988).

DISCUSSION

We have shown that the problem of deciding whether a stimulus is subliminal is difficult. The MAC model provides a reasonable, though not perfect, solution. Draine and Greenwald (1998) and Greenwald, Klinger, and Schuh (1995) advocate a different alternative in which subliminality is not directly assessed. The method works by regressing the priming effect in target identification onto prime detection. The y -intercept of the regression line serves as an estimate of the priming effect when prime detection is at chance. The advantage of this method is that researchers

need not assume at-chance prime detection. Instead, the priming effect for at-chance prime detection is extrapolated from a parametric fit. The method remains controversial, and one of the critiques is that the analyst must assume a functional form between the priming effect and prime detectability (Doshier, 1998; Klauer, Draine, & Greenwald, 1998; Merikle & Reingold, 1998). Our method, in contrast, relies on prime detection alone; there is no need to model the functional relationship between the two performance indices. Consequently, the MAC model may provide more direct evidence for subliminality.

In conclusion, the MAC model stems from the assumption that some stimuli are so impoverished that they are never perceptible. This assumption, coupled with the hierarchical structure, allows for the estimation of the probability that a particular individual is at chance. Perhaps the most attractive feature of the model is that it is not susceptible to the null sensitivity problem. The MAC model provides an incentive to collect more rather than less data, because subliminality can only be claimed with sufficient evidence. We believe that the MAC model will lead to a more precise assessment of the subliminality of priming effects.

AUTHOR NOTE

This research was supported by NSF Grant SES-0351523 and NIMH Grant R01-MH071418. Address correspondence to J. N. Rouder, Department of Psychological Sciences, 210 McAlester Hall, University of Missouri, Columbia, MO 65211 (e-mail: rouderj@missouri.edu).

REFERENCES

- BLACKWELL, H. R. (1953). Psychological thresholds: Experimental studies of methods of measurement. *Bulletin of the Engineering Research Institute of the University of Michigan* (No. 36).
- DAGENBACH, D., CARR, T., & WILHELMSSEN, A. (1989). Task-induced strategies and near-threshold priming: Conscious influences on unconscious perception. *Journal of Memory & Language*, **28**, 412-443.
- DEHAENE, S., NACCACHE, L., LE CLEC'H, G., KOECHLIN, E., MUELLER, M., DEHAENE-LAMBERTZ, G., ET AL. (1998). Imaging unconscious semantic priming. *Nature*, **395**, 597-600.
- DOSHER, B. (1998). The response-window regression method—Some problematic assumptions: Comment on Draine and Greenwald (1998). *Journal of Experimental Psychology: General*, **127**, 311-317.
- DRAINE, S. C., & GREENWALD, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, **127**, 286-303.
- EGAN, J. P. (1975). *Signal detection theory and ROC analysis*. New York: Academic Press.
- GELFAND, A., & SMITH, A. F. M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398-409.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- GREENWALD, A. G., KLINGER, M. R., & LIU, T. J. (1989). Unconscious processing of dichoptically masked words. *Memory & Cognition*, **17**, 35-47.
- GREENWALD, A. G., KLINGER, M. R., & SCHUH, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: Dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General*, **124**, 22-42.
- JEFFREYS, H. (1982). *Theory of probability*. New York: Oxford University Press.
- KLAUER, K. C., DRAINE, S. C., & GREENWALD, A. G. (1998). An unbiased errors-in-variables approach to detecting unconscious cognition. *British Journal of Mathematical & Statistical Psychology*, **51**, 253-267.
- LEE, M., & WAGENMAKERS, E. J. (2005). Bayesian statistical inference in psychology: Comment on Trafimow (2003). *Psychological Review*, **112**, 662-668.

- LU, J., SPECKMAN, P. L., SUN, D., ROUDER, J. N., & MOREY, R. D. (2007). *A Bayesian approach for assessing conscious and unconscious processes in human memory*. Manuscript submitted for publication.
- LUCE, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review*, *70*, 61-79.
- MERIKLE, P. M., & REINGOLD, E. M. (1998). On demonstrating unconscious perception: Comment on Draine and Greenwald (1998). *Journal of Experimental Psychology: General*, *127*, 304-310.
- MURPHY, S. T., & ZAJONC, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality & Social Psychology*, *64*, 723-739.
- NUNNALLY, J. C., & BERNSTEIN, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- PRATTE, M. S. (2007). *Subliminal priming as a task-characteristic artifact*. Unpublished master's thesis, University of Missouri, Columbia.
- REINGOLD, E., & MERIKLE, P. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44*, 563-575.
- ROUDER, J. N., & LU, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review*, *12*, 573-604.
- ROUDER, J. N., LU, J., MOREY, R. D., SUN, D., & SPECKMAN, P. L. (in press). A hierarchical process dissociation model. *Journal of Experimental Psychology: General*.
- ROUDER, J. N., LU, J., SPECKMAN, P. L., SUN, D., & JIANG, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, *12*, 195-223.
- ROUDER, J. N., LU, J., SUN, D., SPECKMAN, P. L., MOREY, R. D., & NAVEH-BENJAMIN, M. (in press). Signal detection models with random participant and item effects. *Psychometrika*.
- ROUDER, J. N., SUN, D., SPECKMAN, P. L., LU, J., & DZHOU, D. (2003). A hierarchical Bayesian statistical framework for response time distributions. *Psychometrika*, *68*, 587-604.
- VAN OPSTAL, F., REYNVOET, B., & VERGUTS, T. (2005). How to trigger elaborate processing? A comment on Kunde, Kiesel, and Hoffmann (2003). *Cognition*, *97*, 89-97.

NOTE

1. A *t* test for this comparison is equivalent to a Bayesian analysis with noninformative priors (Jeffreys, 1982; see Rouder & Lu, 2005 for a tutorial).

APPENDIX

In this appendix, we provide full specification and methods of analysis for the MAC model. Further details can be found in a supplement at www.missouri.edu/~pcl/papers.

Priors

It is necessary to place prior distributions on parameters μ and σ^2 . We use the following:

$$\mu \sim \text{normal}(0, \sigma_0^2),$$

$$\sigma \sim \text{uniform}(0, 1).$$

The value of σ_0^2 must be chosen beforehand. We have performed a number of simulation experiments to understand the effects of this choice. We recommend $\sigma_0^2 = 1$ as this choice leads to conservative estimates of $Pr[x_i < 0 | \text{data}]$. The choice of a uniform on σ is principled. The value of 1.0 is a reasonable upper limit. Consider the case that $\mu = 0$. The corresponding distribution on p_i for those participants above chance is flat when $\sigma = 1$. If $\sigma > 1$, then the distribution of p_i is bimodal, with modes at .5 and 1.0. We consider this unreasonable and exclude it a priori through the upper limit of 1.0 on σ . Other than this exclusion, there is no a priori commitment to any value of σ . These priors are proper, and the propriety of the priors guarantees the propriety of the posteriors (see Rouder & Lu, 2005).

Analysis

The target of analysis in Bayesian statistics is the derivation of the marginal posterior distribution for each parameter. We follow the common approach of deriving closed-form expressions for conditional posterior distributions and sampling these numerically with Markov chain Monte Carlo (MCMC) techniques (Gelfand & Smith, 1990). The derivation of conditional posterior distributions for this application and a discussion of sampling from them may be found in the supplement (www.missouri.edu/~pcl/papers).

We sampled chains of 50,000 iterations with the first 5,000 iterations serving as a burn-in period. The top panels of Figure A1 show chains for select parameters. To keep the plots uncluttered, chains are displayed for the first 5,000 iterations after burn in. Plots for other iterations are highly similar. The bottom panels show the degree of autocorrelation in the chains for the select parameters. As can be seen, mixing is sufficient.

APPENDIX (Continued)

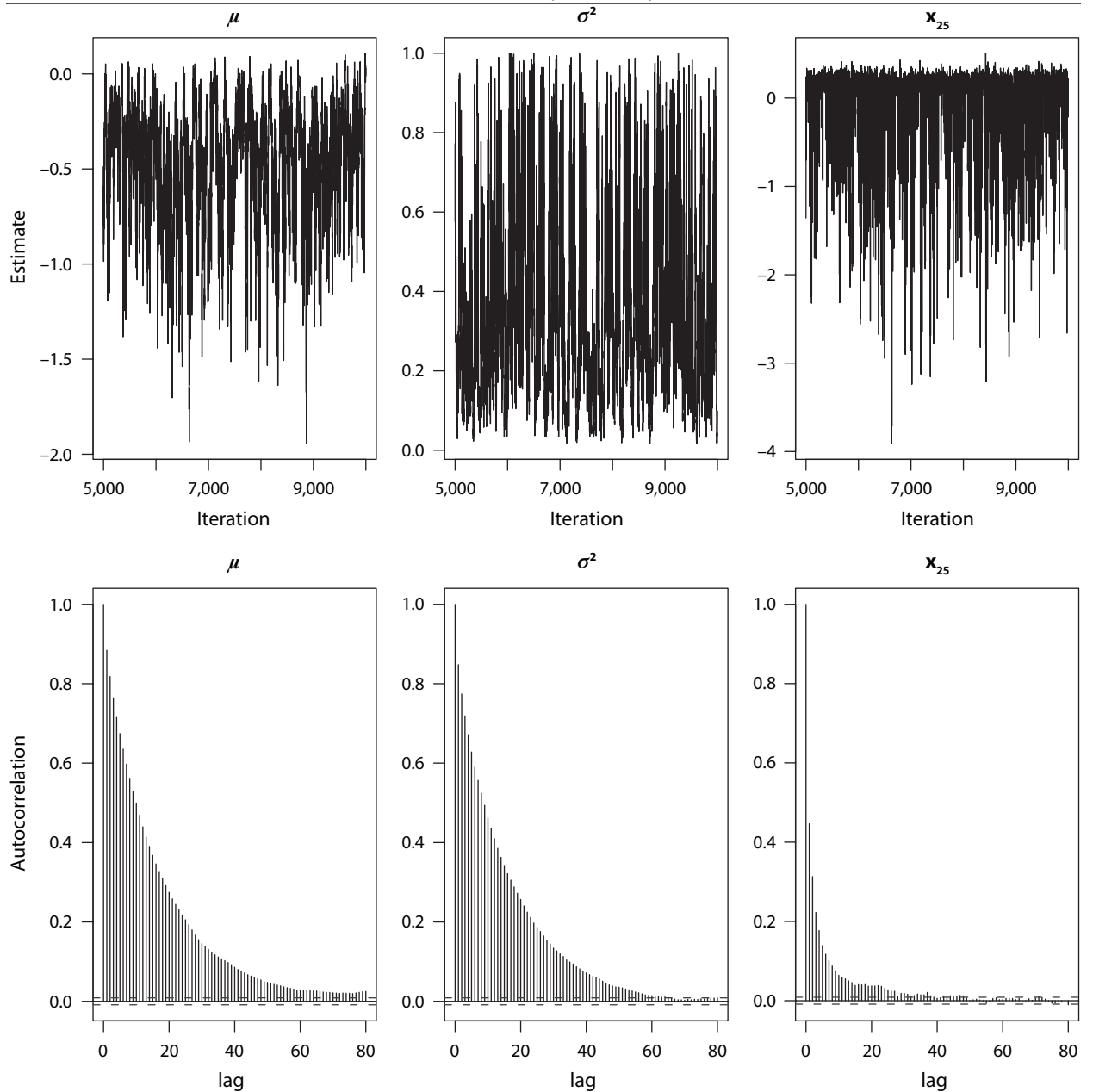


Figure A1. Convergence in MCMC sampling. Top: Values on iterations 5,001 to 10,000 for selected parameters (from left to right: μ , σ^2 , x_{25}). Bottom: Autocorrelation of the chains for the selected parameters. The degree of autocorrelation is offset by the long length of the chains.

(Manuscript received September 12, 2006;
revision accepted for publication October 25, 2006.)