

# Optimal Integration of Auditory and Vibrotactile Information for Judgments of Temporal Order

Ian Ley  
City University

Patrick Haggard  
University College London

Kielan Yarrow  
City University

Recent research that assessed spatial judgments about multisensory stimuli suggests that humans integrate multisensory inputs in a statistically optimal manner by weighting each input by its normalized reciprocal variance. Is integration similarly optimal when humans judge the temporal properties of bimodal stimuli? Twenty-four participants performed temporal order judgments (TOJs) about 2 spatially separated stimuli. Stimuli were auditory, vibrotactile, or both. The temporal profiles of vibrotactile stimuli were manipulated to produce 3 levels of precision for TOJs. In bimodal conditions, the asynchrony between the 2 unimodal stimuli that comprised a bimodal stimulus was manipulated to determine the weight given to touch. Bimodal performance on 2 measures—judgment uncertainty and tactile weight—was predicted with unimodal data. A model relying exclusively on audition was rejected on the basis of both measures. A second model that selected the best input on each trial did not predict the reduced judgment uncertainty observed in bimodal trials. Only the optimal maximum-likelihood-estimation model predicted both judgment uncertainties and weights; the model's validity is extended to TOJs. Alternatives for modeling the process of event sequencing based on integrated multisensory inputs are discussed.

*Keywords:* multisensory integration, temporal order judgment, audition, touch, statistical optimality

## Integrating Cues for Sensory Judgments

In daily life we are often confronted with multiple redundant forms of sensory information that can inform a single perceptual decision. Perhaps the best known example comes from the common situation in which different sensory modalities, such as vision and audition, each provide information about some property of an object (e.g., the location of a loudspeaker). The observer may use one or more sensory channels to locate the loudspeaker, and the question arises as to how information from different modalities is combined or disregarded in order to form a single fused percept (Ernst & Bühlhoff, 2004).

In the natural environment, all sensory modalities that provide relevant information for a particular decision tend to be consistent; that is, each modality specifies the same (true) value for the property that is being judged (although resulting estimates may differ as a result of sensory noise). Under these circumstances, it may be difficult to isolate the contribution of each modality to the final judgment. A common experimental approach has therefore been to introduce a disparity between two sensory inputs (usually

without alerting participants to this fact) and require a judgment about the combined percept. In this way it is possible to dissect the influence of each input by determining how closely the percept follows one input relative to another. For example, Rock and Victor (1964) used a distorting lens to introduce a discrepancy between the seen and felt shape of objects. Their participants reported shapes more consistent with the visual than with the haptic stimulus. This finding was interpreted as showing total visual dominance, although some influence of the haptic stimulus could still in fact be discerned (Ernst & Bühlhoff, 2004). With this finding in mind, it makes sense to talk in terms of the relative weight given to each sensory input, rather than complete reliance on just one source. Another well-studied example in which a sensory discrepancy may go unnoticed is the ventriloquist illusion, in which the ventriloquist's speech appears to come from the mouth of the puppet (Müller, 1838; cited in Bertelson & de Gelder, 2004). In laboratory situations, auditory stimuli are generally mislocalized toward synchronized visual stimuli, with a much smaller influence of auditory stimuli on visual spatial localization (Bertelson & Radeau, 1981; Pick, Warren, & Hay, 1969; but see Alais & Burr, 2004).

Data from experiments investigating spatial judgments have tended to indicate that vision is weighted more heavily than either audition or haptics in bimodal judgments, but this finding does not necessarily generalize to other kinds of judgments. For example, when participants are required to judge the temporal properties of a bimodal stimulus, audition often is found to dominate over vision and touch. Two examples will serve to illustrate this point.

---

Ian Ley and Kielan Yarrow, Department of Psychology, City University, London, England; Patrick Haggard, Institute of Cognitive Neuroscience and Department of Psychology, University College of London, London, England.

Correspondence concerning this article should be addressed to Kielan Yarrow, Social Science Building, City University, Northampton Square, London C1V 0HB, England. E-mail: kielan.yarrow.1@city.ac.uk

First, Shams, Kamitani, and Shimojo (2000, 2002) presented participants with a single brief visual stimulus accompanied by one or more brief auditory stimuli and found that multiple sounds induced participants to report seeing multiple illusory flashes. A similar phenomenon is evident in the tendencies both to perceive the rate of a rapidly fluttering bimodal rhythmic stimulus to be that of the auditory, rather than the visual, component (Recanzone, 2003; Welch, DuttonHurt, & Warren, 1986) and to misperceive the number of taps to the skin in line with the number of accompanying auditory beeps (Bresciani et al., 2005).

Second, building on work by Scheier, Nijhawan, and Shimojo (1999), Morein-Zamir, Soto-Faraco, and Kingstone (2003) had participants perform a temporal order judgment (TOJ) task to determine which of two near-synchronous lights occurred first. When two task-irrelevant sounds were presented, the first one coming just before the first light and the second one coming just after the second light, performance in the TOJ task improved. This “temporal ventriloquism” has been interpreted as a tendency for each sound to attract the temporally closest light and thus to increase the perceived separation between the two lights (although the second sound appears to be a far more powerful attractor; Morein-Zamir et al., 2003).

The classical literature on multisensory integration has given rise to a well-developed theoretical framework for predicting integration effects. One key concept is that the more appropriate a sensory modality is for the particular judgment that is to be made, the more weight that modality receives when the judgment is being reached (Welch & Warren, 1980). To avoid circular reasoning, one must measure appropriateness independently from the bimodal situation. Most explanations refer to the unimodal acuity or precision for the perceptual property that is being assessed (e.g., Welch, Widawski, Harrington, & Warren, 1979; see Welch & Warren, 1980, for a review of the early literature). This is entirely in accord with the previously reviewed studies showing a visual advantage for spatial judgments and an auditory advantage for temporal judgments. Although comparisons are complicated by the differences between stimuli used in different sensory modalities, the visual system generally displays excellent spatial acuity but exhibits low-pass temporal filtering characteristics that may eliminate fine temporal detail (e.g., Hawken, Shapley, & Grosop, 1996). Audition shows the opposite pattern, and tactile acuity falls between these two extremes. Compare, for example, the threshold for detecting a high-amplitude sinusoidal modulation of light intensity (around 50 Hz; e.g., de Lange, 1958) with the thresholds for detecting such modulation for broadband vibrotactile or auditory noise, which lie at around 300 and 1000 Hz, respectively (Viemeister, 1979; Weisenberger, 1986).

### Statistical Optimality and Cue Integration

Psychological concepts such as modality appropriateness and precision may be difficult to assess in a mathematically rigorous manner. Recently, a number of researchers have attempted to formally describe the manner in which different inputs are weighted and combined for sensory judgments (Ernst & Bühlhoff, 2004). A statistical framework has been used to define the optimal method for combining inputs (with optimality defined in terms of maximizing precision), and psychophysical performance has been compared with this prediction. It is assumed that sensory inputs

will contain varying degrees of environmental noise and will be further contaminated by noise during sensory transduction and transmission. Hence, the brain is faced with noisy information and will produce sensory estimates about the state of the environment that vary from trial to trial. In general, any sensory estimate that is not subject to systematic bias will be correct on average but will display variance. Under these conditions, with the assumption of independent Gaussian noise, formal analysis shows that the combined sensory estimate ( $\hat{S}$ ) is best (i.e., has the lowest variance) when all sensory estimates ( $\hat{S}_i$ , with the subscript referring to the sensory modality) are combined in a weighted average (Equation 1). The weight given to a specific estimate ( $w_i$ ) should be inversely proportional to the variance of that sensory estimate ( $\sigma_i^2$ ; Equation 2).

$$\hat{S} = \sum_i w_i \hat{S}_i \text{ with } \sum_i w_i = 1 \quad (1)$$

$$w_j = \frac{1/\sigma_j^2}{\sum_i 1/\sigma_i^2} \quad (2)$$

Individual sensory variability can be estimated directly by recording estimates across a number of trials in unimodal conditions. If we concentrate on the bimodal case (e.g., combining an auditory stimulus with variance  $\sigma_A^2$  and a tactile stimulus with variance  $\sigma_T^2$ ), we can rewrite Equation 2 as follows:

$$w_T = \frac{1/\sigma_T^2}{(1/\sigma_T^2) + (1/\sigma_A^2)} = \sigma_A^2 / (\sigma_A^2 + \sigma_T^2) \quad (3)$$

When the optimal weighting rule is followed, variability in the multimodal case will always be lower than the variability in the best of the contributing unimodal sensory estimates. The bimodal variance  $\sigma_{AB}^2$  can be determined from Equation 4:

$$\sigma_{TA}^2 = \sigma_A^2 \sigma_T^2 / (\sigma_A^2 + \sigma_T^2) \leq \min(\sigma_A^2, \sigma_T^2) \quad (4)$$

Experiments can be devised to test how closely human judgments conform to the predictions of such an optimal maximum likelihood estimation (MLE) model. Conformity suggests that the dominance of one sensory modality for any particular judgment does not imply an inflexible reliance on that modality for all judgments of a particular type (e.g., spatial, temporal) but rather a tendency to favor more precise inputs over less precise inputs, regardless of their sensory origin. Examples of this approach include the work of Hillis, Watt, Landy, and Banks (2004) for combining of depth cues; Van Beers, Sittig, and Gon (1999) for combining visual and proprioceptive information in two dimensions to estimate the position of an unseen arm; Heron, Whitaker, and McGraw (2004) for estimating the effect of a transitory auditory stimulus on the perceived position of reversal of a moving visual stimulus; and Shams, Ma, and Beierholm (2005) for assessing the impact of transient beeps on the perceived number of transient flashes and vice versa. The current research was influenced heavily by the experimental design used by Ernst and Banks (2002), so that study will be described in more detail as a concrete example.

Ernst and Banks (2002) used two force-feedback devices attached to the forefinger and thumb to create virtual haptic objects (raised horizontal bars). Spatially compatible visual stimuli were created with random dot stereograms. In unimodal conditions, participants had to compare a stimulus of standard (55-mm) height

with a comparison stimulus that varied in height about this standard value. Noise could be added to the visual stimulus in order to vary the difficulty of the discrimination. Variance was estimated as the squared width of the psychometric function fitted to a participant's responses. Stimuli were designed so that the haptic stimulus variance fell within the range of variances for the visual stimuli. In bimodal conditions, the different visual stimuli were combined with the haptic stimulus in the same judgment. A further manipulation varied each component of the bimodal standard, so that it averaged 55 mm but might, for example, be composed of a shorter visual stimulus and a longer haptic stimulus. This manipulation allowed the authors to determine the weight given to each modality by determining the change in the point of subjective equality produced by a particular visual-haptic discrepancy. Ernst and Banks (2002) found that the variances and weights determined from the bimodal condition were in close agreement with the optimal MLE model described above. Visual weights decreased as the visual stimulus became noisier, and bimodal variance was always lower than the better of the two contributing unimodal variances. A very similar approach was used by Alais and Burr (2004), who varied the spatial precision of a visual stimulus to show that the ventriloquist effect represented an example of statistically optimal integration.

### Rationale for the Current Investigation Assessing TOJs

To date, those multisensory studies that have tested the optimal integration model have assessed mainly spatial judgments, in which vision would traditionally be assumed to dominate. More recently, studies have begun to assess judgments that may partially depend upon temporal acuity, such as counting the number of rapid transient events in a short train (Bresciani, Dammeier, & Ernst, 2006; Bresciani & Ernst, 2007; Shams et al., 2005; Wozny, Beierholm, & Shams, 2008) or comparing the rate of a flickering or fluttering stimulus (Roach, Heron, & McGraw, 2006; Wada, Kitagawa, & Noguchi, 2003). These studies have generally supported the MLE model, albeit with the addition of model parameters intended to downplay integration in the context of obvious sensory discrepancies. However, tasks like these do not really investigate timing mechanisms but rather mechanisms of numerosity coding and rate perception. In the current study, we wished to assess whether judgments about the sequencing of events in time would also exhibit statistically optimal multisensory integration. For example, when we witness a car crash and must place it in temporal context, do we perceive the time of collision based on either visual or auditory information or do we combine both types of information in an optimal manner? To this end, we developed a TOJ task in which variability could be manipulated in a manner analogous to the changes in visual noise introduced by Ernst and Banks (2002).

In TOJ tasks, participants are presented with two brief stimuli in rapid temporal succession and are asked to discriminate which came first. The method has been widely used throughout the history of experimental psychology (Spence, Shore, & Klein, 2001). It provides the most direct way to assess how events are ordered to construct a subjective timeline. Studies on temporal ventriloquism have indicated that multisensory integration can be assessed with TOJs (Morein-Zamir et al., 2003; Scheier et al., 1999). Furthermore, irrelevant stimuli in a second modality that

are presented either in identical or opposite order to the attended stimuli in a TOJ task have been shown to affect judgments of temporal order (Kitazawa et al., 2007; Sherrick, 1976). However, we are not aware of any previous attempt to compare precisely the TOJs made in a bisensory task with predictions based on Bayesian models. Given the precision of these models, this represents an important gap in our knowledge. To investigate the issue, we presented vibrotactile and auditory stimuli, presented on the left and right sides, in unisensory and bisensory conditions while we manipulated the difficulty of the vibrotactile judgment. On some bisensory trials, we introduced a very small asynchrony between the auditory and vibrotactile components of the right-hand stimulus. Our participants judged the side from which the first stimulus came. In this way we were able to both generate an experimental situation that promoted optimal integration and accurately test the optimal MLE model against judgments of temporal order.

To evaluate the MLE model, we generated predictions from three models and attempted to demonstrate statistically significant mean deviations from two of the three in order to demonstrate the plausibility of the remaining model. We also compared the models directly on their squared errors of prediction. The first model assumes that because audition has generally been found to dominate over both vision and touch for temporal judgments, participants would simply rely on the auditory information and ignore the vibrotactile information. We refer to it as the rely on audition model. The second model assumes that participants would use only one sensory modality on each trial but that they would flexibly select the sensory modality depending on which modality contained the stimulus that provided the most precise information about temporal order. We refer to this as the best sensory estimate model. Finally, the third model assumes statistically optimal bisensory integration. We refer to it as the MLE model.

## Method

### *Participants*

Initially, 24 participants were tested. On the basis of this data set it was impossible to discriminate the MLE model from the best sensory estimate model, as neither was statistically distinguishable from the data. In order to better discriminate between models, we decided to reject and replace a subset of outlying participants to make our sample more homogeneous. Participants who returned any point of subjective simultaneity or judgment uncertainty estimate (see *Analysis*, below) that deviated by more than four standard deviations from the group mean in 1 or more of the 13 unimodal and bimodal conditions were excluded. This led to the rejection of 4 participants, who were then replaced to yield a final sample of 24 participants (13 men; mean age = 25.6 years,  $SD = 3.9$ ).

### *Apparatus and Stimuli*

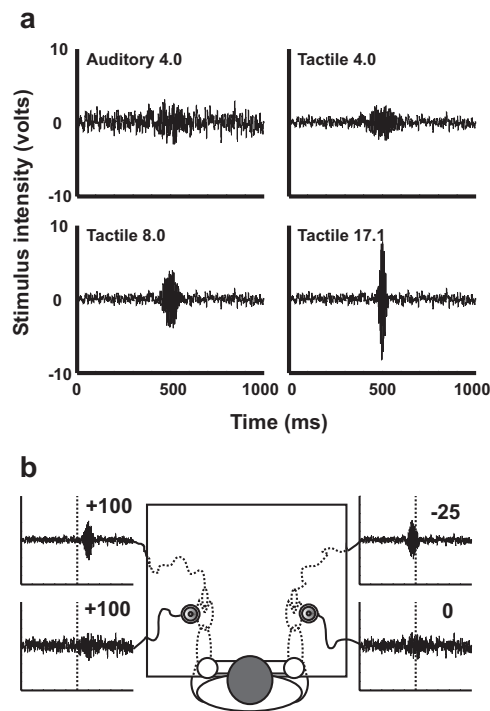
The experiment was controlled by a PC that produced auditory and vibrotactile stimuli at 44100 Hz with a 12-bit A/D card (National Instruments DAQCard 6715, Austin, TX). We confirmed the correct timing of output signals using a 20-MHz storage oscilloscope (Gould DSO 1604, Valley View, OH). Both auditory and vibrotactile stimuli were Gaussian-windowed 120-Hz sinu-

soids. The Gaussian windowing procedure was used to produce stimuli that could be temporally smeared to varying extents; a small standard deviation produced a brief, sharply defined stimulus, whereas a large standard deviation produced a longer stimulus with an indistinct peak. In addition to varying the standard deviation of the window, we varied its peak intensity. The area under the Gaussian window was held constant, such that a stimulus with a small standard deviation had a high peak intensity and vice versa. The resulting windowed sinusoids were then embedded in low-frequency background noise. A 1-s segment of white noise was produced by generating random voltages from a uniform distribution. This signal was digitally filtered with a second-order bidirectional Butterworth filter with a high cut frequency of 240 Hz. Windowed sinusoids were added to low-pass noise to produce our final stimuli. We used three vibrotactile stimuli and one auditory stimulus, with parameters selected based on pilot work and modified slightly after the first 8 participants had been analyzed. For the first 8 participants, the auditory stimulus had a maximum peak-to-peak voltage of 4.0 V and a standard deviation of 59 ms, and it was embedded in filtered noise with a root-mean-square voltage of 0.785 V. The three vibrotactile stimuli had maximum peak-to-peak voltages of 4.0, 8.0, and 17.1 V, with standard deviations of 59, 29, and 13 ms, respectively. They were embedded in filtered noise with a root-mean-square voltage of 0.392 V. For the remaining participants, both the auditory stimulus and the intermediate vibrotactile stimulus were adjusted to have a maximum peak-to-peak voltage of 6.5 V and a standard deviation of 36 ms. Background filtered noise was unchanged, being higher in the auditory case. These changes were made in order to provide a better match between the intermediate vibrotactile stimulus and the auditory stimulus (and also a greater difference between these stimuli and the two extreme vibrotactile stimuli), which maximizes the difference in model predictions (see below). The stimuli used for the first 8 participants are shown in Figure 1a. Note that these are the voltages sent to the vibrotactile and auditory actuators and not necessarily the physical stimuli that were produced by these components.

Participants sat with their head on a chin rest; its height had been adjusted to suit their preferred posture. Auditory stimuli were presented from two small speakers, located below and in front of the participant's head on a desktop, one to the left and one to the right. The speakers were spaced 30 cm apart and were 20 cm in front of the headrest. Their distance below ear level varied from participant to participant but typically was around 40 cm. Vibrotactile stimuli were delivered via two small (1-cm-diameter) ceramic piezoelectric discs coated in plastic. The discs were driven from a custom-built amplifier and did not produce audible noises with any of the stimuli we used. They were attached to the underside of the desk in front of the participant, with one disc mounted directly below each auditory speaker. The discs were gripped comfortably between index finger and thumb.

### Design

The experiment consisted of two phases: a unimodal phase, in which baseline auditory and vibrotactile performance were assessed separately, and a bimodal phase in which combined auditory/vibrotactile stimuli were presented. The order of the two phases was counterbalanced across participants. In the unimodal



*Figure 1.* Schematic of experimental stimuli and methods. a: Stimuli sent to auditory and vibrotactile actuators for the first 8 participants. Gaussian-windowed sine waves were embedded in low-frequency noise. The width of the Gaussian window and the peak intensity of the signal were manipulated to adjust discriminability. b: The approximate position of the actuators is shown alongside stimuli presented on an example trial from the bimodal phase of the experiment. Dashed lines indicate objects beneath the desktop. On this trial, the combined left-hand stimulus is presented 100 ms after the right-hand stimulus, which itself contains a small (25-ms) discrepancy between the unimodal auditory (bottom) and vibrotactile (top) components. We used this discrepancy to assess the weight given to vibrotactile in the overall judgment.

phase, participants received one block of 75 trials with auditory stimuli and one block of 225 trials with vibrotactile stimuli. The auditory block contained a single kind of stimulus. The vibrotactile block contained 75 trials with each of the three types of vibrotactile stimulus, in a pseudorandom order. The order in which the two unimodal blocks that made up the unimodal phase were received was counterbalanced across participants.

In the bimodal phase, participants received a single block of 675 trials. A two-factor ( $3 \times 3$ ) design was used. The first factor, tactile difficulty, varied the parameters of the tactile stimulus, as discussed above. The second factor, auditory–tactile disparity, varied the temporal position of the peak of the right-sided vibrotactile stimulus relative to the peak of the right-sided auditory stimulus. The peaks of the two left-sided stimuli always coincided exactly. Hence on each trial the auditory stimulus was combined with one of three tactile stimuli, with the right-sided tactile stimulus presented at one of three temporal offsets relative to the right-sided auditory stimulus ( $-25$  ms,  $0$  ms, and  $25$  ms). An example of one trial from the bimodal phase of the experiment is shown in Figure 1b. Participants received 75 trials from each condition in a pseudorandom order.

### Procedure

In the unimodal phase, one stimulus was delivered to the right side and one to the left side on each trial. Both stimuli came from the same modality and were generated using the same Gaussian window (i.e., had the same temporal smear and peak intensity). The noise components of each of the two 1-s-long stimuli were identical, and they began and finished at the same time. In contrast, the windowed sinusoidal components of the two stimuli could be temporally offset from one another. The Gaussian window for the right-sided stimulus peaked at 500 ms, exactly halfway into the stimulus. The Gaussian window for the left-sided stimulus could peak anywhere from 300 ms before the right-sided stimulus to 300 ms after the right-sided stimulus. Participants judged whether the left or the right stimulus had occurred first; they had been directed to attend to the distinct sinusoidal peaks, not the onset of the background noise. Their responses were entered into the computer by the experimenter.

The delay between the right- and left-sided stimuli varied from trial to trial. It was selected randomly on each trial from a condition-specific distribution. There was a single distribution for the auditory block and separate ones for each of the three stimuli used in the vibrotactile block. Each distribution was initially uniform and contained delay values from  $-140$  to  $140$  ms in 20-ms increments, but it was updated after each accepted trial according to the generalized P'olya urn model (Rosenberger & Grill, 1997;  $k = 8$ ). Distributions could therefore expand to include delay values from  $-300$  to  $300$  ms. This procedure produces many values close to the point of subjective simultaneity.

In the bimodal phase, the procedure was similar, but four stimuli, two vibrotactile and two auditory, were delivered on each trial. Participants were told that the two stimuli on the left would peak at the same time, as would the two stimuli on the right, and were required to judge which of the two combined (bimodal) stimuli came first. In fact, while the peaks of the two stimuli on the left were indeed synchronous, either the two stimuli on the right could have synchronous peaks or the vibrotactile component could be offset slightly from the auditory component by 25 ms in either direction. The auditory component of the right-sided stimulus peaked at 500 ms. The vibrotactile component of the right-sided stimulus peaked at 475, 500, or 525 ms. The two left-sided stimuli peaked with a delay of  $-300$  to  $300$  ms that was determined relative to the auditory component of the right-sided stimulus. Separate adaptive distributions were maintained to randomly select delays in each of the nine conditions (three levels of tactile difficulty crossed with three levels of right-sided auditory–tactile disparity).

### Analysis

The proportion of times that a participant judged the right-sided stimulus to have occurred first for each delay value that had been presented was determined separately in each condition. Cumulative Gaussian psychometric functions were fitted to these data using the *psignifit* toolbox Version 2.5.6 for Matlab (see <http://bootstrap-software.org/psignifit/>), which implements the maximum-likelihood method described by Wichmann and Hill (2001). Points of subjective simultaneity (PSSs) and judgment uncertainty were estimated from these functions. A PSS was

estimated from the delay value at which the “right first” judgment occurred with a probability of 0.5. Judgment uncertainty is a threshold value similar to the commonly assessed just-noticeable difference (JND), and it was estimated as the difference between the delay values that yielded “right first” judgments with probabilities of 0.5 and 0.84.

The bimodal conditions were used to determine the manner in which participants combined information from the tactile and auditory modalities to reach a temporal judgment. We wished to assess both the bimodal judgment uncertainty and the weight given to the tactile modality for the combination of the auditory stimulus with each of the three different vibrotactile stimuli. To determine the bimodal judgment uncertainty, we averaged the judgment uncertainty values estimated for the three bimodal conditions sharing a particular tactile stimulus (e.g., the  $-25$ -ms, 0-ms, and 25-ms disparity conditions with the tactile stimulus of gradual temporal profile), as the manipulation of auditory–tactile asynchrony for the right-hand stimulus would not be expected to change the slope of the psychometric function, given the range of stimuli we used. To determine the weight given to the tactile modality, we determined the difference between the PSSs estimated for the 25 and  $-25$  ms disparity conditions. To normalize this value, we divided it by 50, because 50 is the expected change in PSS if participants had based their judgments entirely on the tactile modality.

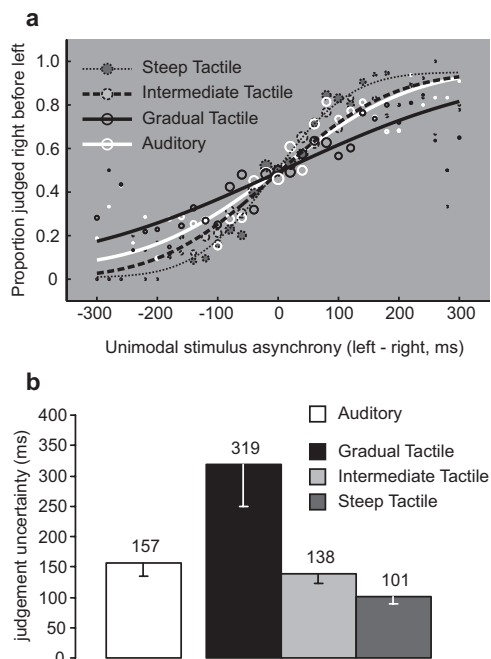
Data from the three unimodal conditions were used to predict performance in the bimodal condition under each of the three models considered in the Introduction. For the rely on audition model, predicted judgment uncertainty in each of the three bimodal cases was equal to judgment uncertainty in the unimodal auditory condition, and the predicted tactile weights were all zero. For the best sensory estimate model, the predicted judgment uncertainty was taken as the lower of the unimodal auditory or the unimodal tactile judgment uncertainties for the relevant tactile stimulus profile. Predicted tactile weights were set to zero if the auditory judgment uncertainty was lower than the relevant tactile judgment uncertainty and to 1.0 if the opposite was the case. Finally, for the MLE model, predicted judgment uncertainties and tactile weights were determined according to Equations 4 and 3 in the Introduction, respectively.

Traditional measures of goodness of fit were not appropriate to test our models, because we were not fitting models with free parameters to a data set but rather were testing the predictions of each model on a new data set. The predictions from each of the three models were therefore compared with the values estimated directly from the bimodal conditions. Predictions also were used to produce an additional measure of predictive success: mean squared deviation. All comparisons were made using standard parametric statistics (analyses of variance [ANOVAs] with Greenhouse–Geisser corrections for violations of sphericity and repeated measures  $t$  tests).

## Results

### Unimodal Data

Figure 2 shows the results of the four unimodal conditions in which participants performed TOJs between two lateralized stimuli, both coming from the same sensory modality. A single audi-



**Figure 2.** Judgment uncertainty in unimodal conditions. a: Illustration of sigmoid fitting, based on combined data from all 24 participants. The size of each data point provides a rough guide to the number of observations collected. Note that because an adaptive procedure was used to determine the range of asynchronies each participant received, the more extreme asynchronies were delivered only to more uncertain participants, and then only rarely, whereas the central asynchronies reflect judgments from all participants. This explains the apparent rise in uncertainty at the extremes of the graph. b: Mean judgment uncertainties across participants, determined by individual fits to each participant's data. Error bars denote standard errors.

tory stimulus was assessed, along with three vibrotactile stimuli constructed so as to vary the difficulty of this temporal discrimination. Figure 2a illustrates how judgment uncertainty was determined. To produce this figure, we combined the data from all 24 participants in each unimodal condition and fitted the data with a cumulative Gaussian sigmoid function. The slopes of the fitted sigmoids reflect the difficulty of each judgment, with steep slopes indicating less noisy judgments. Figure 2b shows the mean judgment uncertainty in each condition determined from our actual analysis. To generate these data, we fitted a different sigmoid to each individual participant's data in each condition, calculated judgment uncertainty, and then averaged across the group. Hence judgment uncertainties differ somewhat from those that would be estimated from the fits shown in Figure 2a. Figure 2b indicates that our manipulation of difficulty in the vibrotactile conditions was successful. The tactile stimulus with the gradual temporal profile yielded the highest judgment uncertainty, which was considerably higher than that of the auditory condition; the steeply profiled tactile condition yielded the lowest judgment uncertainty, which was lower than that of the auditory condition. The intermediate tactile condition was intermediate in difficulty between these extremes and showed judgment uncertainty similar to but slightly lower than that of the auditory condition. Although our measure of judgment uncertainty is somewhat unusual, being the difference

between the midpoint of the psychometric function and the 0.84 point (one standard deviation along the cumulative Gaussian), the judgment uncertainties can be easily converted to JNDs (the difference between the 0.75 and 0.5 points) to facilitate comparison with other studies. The JND is 0.67 times the judgment uncertainty value we report; for the unimodal data, JNDs therefore ranged from 214 ms (for the gradual tactile stimulus) to 68 ms (for the steep tactile stimulus).

The range of uncertainties in our data is not quite as great as would have been ideal, but differences across tactile conditions were statistically reliable as assessed with a one-way repeated-measures ANOVA,  $F(1.24, 28.43) = 8.31, p = .005$ . This indicated that our experimental manipulations could provide a suitable test of our three models if we used data collected in bimodal conditions.

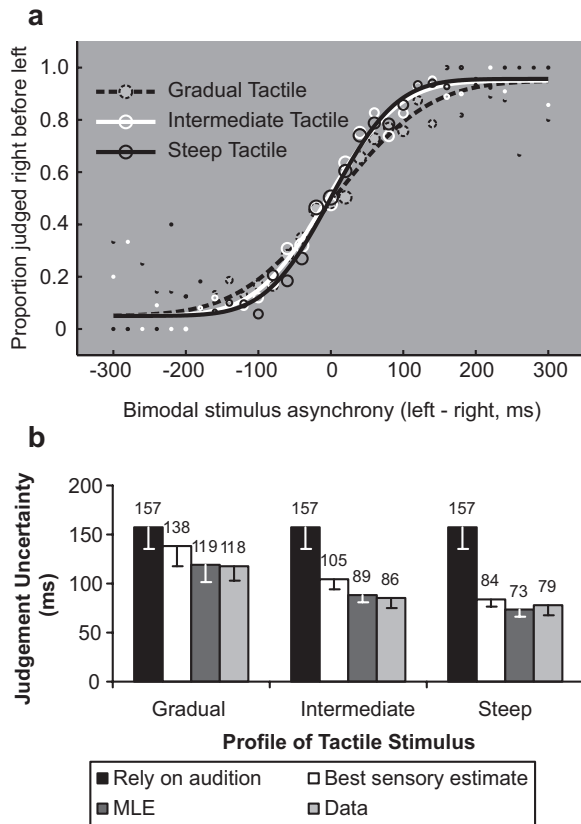
### *Bimodal Data: Judgment Uncertainty*

In addition, each participant discriminated the temporal order of combined bimodal stimuli presented to the left and right. The auditory stimulus could be combined with each of the three vibrotactile stimuli to produce these bimodal stimuli. Figure 3a illustrates how judgment uncertainty was determined. The combined data across all 24 participants are displayed for the three conditions that combined the auditory stimulus with each of the vibrotactile stimuli without introducing any between-modality asynchrony. The best fitting sigmoids are shown also.

For our actual analysis, we fitted each participant's data in each condition with a separate sigmoid to determine judgment uncertainty. We then produced a single estimate of judgment uncertainty for each participant and each combination of the auditory stimulus with one of the three vibrotactile stimuli. To do this, we averaged across the three conditions that introduced a slight between-modality asynchrony, as any differences between these would be relevant only for assessing weights (see below). In addition, we used each participant's judgment uncertainties in the unimodal conditions to predict that participant's bimodal scores based on three possible models: the rely on audition model, which assumes that the vibrotactile input is disregarded; the best sensory estimate model, which assumes that on each trial the more precise modality is used; and the MLE model, which combines the two inputs weighted according to their precision. Figure 3b shows the mean judgment uncertainty averaged across all participants, along with the mean predicted judgment uncertainty for each of the three models. These judgment uncertainties equate to JNDs of 79, 58, and 53 ms for the gradual, intermediate, and steep conditions, respectively.

The pattern of mean judgment uncertainties supports the MLE model. When the vibrotactile component of the bimodal stimulus had a gradual or intermediate temporal profile, performance was almost identical to the MLE prediction and better than either the rely on audition or the best sensory estimate predictions. For the steep vibrotactile stimulus, MLE and best sensory estimate predictions were very close to one another, and actual performance was about midway between these two predictions.

A comparison of the bimodal judgment uncertainties with the predictions of each model, like that shown in Figure 3b, allows us to visualize the difference between the model's predictions and the mean of the data. It is possible, however, for a model to achieve a



**Figure 3.** Comparison of bimodal judgment uncertainty and model predictions for three models of bisensory integration. **a:** Illustration of sigmoid fitting, based on combined data from 24 participants in bimodal conditions without any between-modality asynchrony. The size of each data point provides a rough guide to the number of observations collected. See legend to Figure 2 for an explanation of noise at high absolute asynchronies. **b:** Mean bimodal judgment uncertainties across participants, determined by individual fits to each participant's data, and mean model predictions across participants, based on unimodal judgment uncertainties. Error bars denote standard errors. MLE = maximum likelihood estimation.

good prediction in terms of the mean deviation but to be relatively poor at predicting the data of each individual participant (so long as the different errors of prediction for each participant sum to around zero). For this reason, we also calculated the mean squared deviation of each model from the bimodal data, to measure variation in the model's predictive success. These data are shown in Table 1. Table 1 conforms broadly with Figure 3b and shows that the MLE model had the lowest mean squared error for all three tactile profiles.

To investigate the predictive power of each model statistically, we first carried out a series of two-way ( $2 \times 3$ ) repeated-measures ANOVAs to assess mean deviations of model predictions from the bimodal data. Each ANOVA compared one model's predictions with the empirical data (the factor model) at each level of vibrotactile stimulus temporal profile (the factor tactile profile). Comparing empirical data with the rely on audition model, the ANOVA revealed a main effect of model,  $F(1, 23) = 12.76, p = .002$ ; a main effect of tactile profile,  $F(1.19, 27.43) = 10.42, p = .002$ ;

and an interaction,  $F(1.19, 27.43) = 10.42, p = .002$ . The main effect of model allows us to reject this model as a viable explanation of bisensory performance in our task. The interaction shows that this rejection is most compelling when the vibrotactile stimulus had a steeper profile.

Comparison of the data with the best sensory estimate model revealed a main effect of model,  $F(1, 23) = 4.30, p = .049$ , and a main effect of tactile profile,  $F(1.28, 29.34) = 10.56, p = .002$ , with no interaction,  $F(1.73, 39.85) = 1.30, p = .280$ . Once again, the main effect of model allows us to reject this explanation.

In contrast to the findings for the first two models, comparison of empirical data with the MLE model revealed a main effect of tactile profile,  $F(1.19, 27.45) = 11.36, p = .001$ , but no main effect of model,  $F(1, 23) = 0.003, p = .956$ , and no interaction,  $F(1.73, 39.76) = 0.62, p = .521$ . Of the models we investigated, only this one cannot be rejected on the basis of observed judgment uncertainties in bimodal conditions.

A supplementary analysis was carried out on the mean squared deviations for each model. A two-way ( $3 \times 3$ ) repeated-measures ANOVA compared mean-squared deviations of bimodal judgment uncertainty data from model predictions, with the factor model comparing the three models, and the factor tactile profile comparing the three different vibrotactile stimulus profiles. There was a main effect of model,  $F(1.01, 23.12) = 4.85, p = .038$ , but no main effect of tactile profile,  $F(1.09, 24.97) = 2.56, p = .12$ , and no interaction,  $F(1.08, 24.86) = 3.08, p = .089$ . Pairwise follow-ups (Tukey's least significant difference) investigating the main effect of model collapsed across the three levels of tactile profile showed significant differences between all three models: The MLE model had a significantly lower mean squared deviation than did either the best sensory estimate model ( $p = .025$ ) or the rely on audition model ( $p = .035$ ), and the best sensory estimate model had a significantly lower mean squared deviation than did the rely on audition model ( $p = .042$ ).

During testing, we rejected and replaced four participants in order to homogenize our sample and increase the likelihood of being able to discriminate statistically between them. This change was intended to be neutral with regard to the proportion of participants that appeared to provide support for each model. If a model is accurate, we would expect approximately half of the participants to yield bimodal estimates above model predictions and half of the participants to yield estimates below model predictions. In the initial sample, the proportion of participants who scored below

**Table 1**  
*Mean Squared Deviation (MSD) of Bimodal Judgment Uncertainties From the Predictions of the Models*

Model	Gradual tactile profile		Intermediate tactile profile		Steep tactile profile	
	MSD	SE	MSD	SE	MSD	SE
Rely on audition	6,458	3,045	13,654	5,600	15,554	6,492
Best sensory estimate	3,078	1,043	2,001	566	1,672	711
MLE	1,539	393	1,389	663	1,534	875

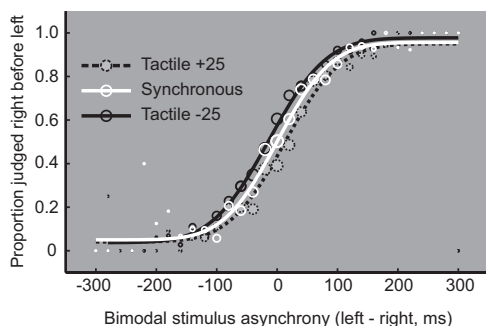
*Note.* Data are shown separately for the combination of the auditory stimulus with each of the three vibrotactile stimuli. MLE = maximum likelihood estimation.

the predictions of the MLE model was 11/24 in the gradual tactile condition, 10/24 in the intermediate condition, and 13/24 in the steep condition. For the final sample, these proportions changed only slightly, to 11/24, 12/24, and 12/24, respectively. Proportions scoring below the predictions of the best sensory estimate model were 15/24, 16/24, and 17/24 for the initial sample and 16/24, 17/24, and 17/24 for the final sample. Finally, for the rely on audition model, initial proportions scoring below predictions were 17/24, 19/24 and 21/24, with proportions in the final sample being 18/24, 20/24 and 22/24 respectively. Model predictions and data are shown for each participant separately in the Appendix, along with 95% confidence intervals.

### Bimodal Data: Vibrotactile Weights

In bimodal conditions, the relationship between the two unimodal stimuli that made up the bimodal stimulus on the right-hand side was manipulated in a subtle manner. The vibrotactile component could be synchronous with the auditory component or could either precede or follow it by 25 ms. Participants were not alerted to this manipulation (they were told that the two components would always be synchronized), and only 1 participant spontaneously asked about such a manipulation. On further discussion it transpired that this question had been motivated by his expert understanding of multisensory research methodologies rather than any perceived asynchrony.

Figure 4 illustrates how changes in PSSs were used to determine the weight given to the vibrotactile modality for each bimodal combination of the auditory and one of the three vibrotactile stimuli. Data are shown for all 24 participants in the three bimodal conditions in which the auditory stimulus was combined with the steeply profiled vibrotactile stimulus. Best fitting sigmoids are displayed for conditions in which the right-hand bimodal stimulus included a small (25-ms) asynchrony between auditory and vibrotactile components or no such asynchrony. The curves are clearly shifted along the horizontal axis, and this indicates that the change in the relative timing of the vibrotactile stimulus influenced the participants' PSSs.



**Figure 4.** Illustration of method for determining tactile weights, based on combined data from 24 participants in bimodal conditions with a steep vibrotactile component stimulus and a  $-25$ -,  $0$ -, or  $25$ -ms asynchrony between the right-hand auditory and vibrotactile component stimuli. The shift in the PSS is used to estimate the weight given to vibrotaction. The size of each data point provides a rough guide to the number of observations collected. See legend to Figure 2 for an explanation of noise at high absolute asynchronies.

For our actual analysis, we fitted each participant's data with a separate sigmoid in each bimodal condition and determined the PSS. We used the change in PSS between the  $-25$ -ms and  $25$ -ms asynchrony conditions for a particular vibrotactile profile to determine the weight given to that vibrotactile stimulus. We used the judgment uncertainties estimated for each participant in unimodal conditions to form predictions about vibrotactile weights based on our three models (rely on audition, best sensory estimate, and MLE). Figure 5 shows mean vibrotactile weights across participants based on bimodal data, along with mean predictions for the three models.

The pattern of vibrotactile weights is not entirely consistent with any of the three models. It is closer to the predictions of the best sensory estimate and MLE models, which both predict that weights should rise as vibrotactile strength increases, than to those of the rely on audition model, which does not, but the slight drop from the medium to high strength conditions is not predicted by any model. However, the large error bars suggest a cautious interpretation. Overall, means are closest to the predictions of the MLE model.

As we did with the judgment uncertainty data, we determined mean squared deviations as an additional measure of each model's predictive success. Mean squared deviations are shown in Table 2. In general, the MLE model produced considerably lower mean squared deviations than the rely on audition model and slightly lower values than the best sensory estimate model, although the best sensory estimate model was most successful for the gradual tactile condition.

To investigate the predictive power of each model statistically, we first carried out three two-way ( $2 \times 3$ ) repeated-measures ANOVAs to assess mean deviations of each model's tactile weight predictions from the weights estimated using the bimodal data. Each ANOVA compared one model's predictions with the empirical data (the factor model) at each level of vibrotactile stimulus temporal profile (the factor tactile profile).

Comparison of the empirical data with the rely on audition data showed there was a main effect of model,  $F(1, 23) = 32.68$ ,  $p < .001$ , but no main effect of tactile profile,  $F(1.57, 36.18) = 0.98$ ,  $p = .367$ , and no interaction,  $F(1.57, 36.18) = 0.98$ ,  $p = .367$ . The main effect of model allows us to reject this model as an explanation of bisensory performance in our TOJ task. However, when we compared empirical data with the best sensory estimate and MLE models, we were unable to reject either model. In both cases there was a main effect of tactile profile, best sensory estimate,  $F(1.52, 35.04) = 4.61$ ,  $p = .025$ ; MLE,  $F(1.48, 34.01) = 3.63$ ,  $p = .05$ , but no main effect of model, best sensory estimate,  $F(1, 23) = 0.45$ ,  $p = .507$ ; MLE,  $F(1, 23) = 0.11$ ,  $p = .740$ , and no interaction, best sensory estimate,  $F(1.89, 43.37) = 1.38$ ,  $p = .262$ ; MLE,  $F(1.72, 39.47) = 0.69$ ,  $p = .485$ .

A supplementary analysis was carried out on the mean squared deviations for each model. A two-way ( $3 \times 3$ ) repeated-measures ANOVA compared mean-squared deviations of bimodal tactile weight estimates from model predictions, comparing across the three models, and also the three different vibrotactile stimulus profiles. There was a main effect of model,  $F(1.07, 24.69) = 5.79$ ,  $p = .022$ , but no main effect of tactile profile,  $F(1.35, 31.03) = 2.01$ ,  $p = .163$ , and no interaction,  $F(2.32, 53.35) = 0.77$ ,  $p = .486$ . Pairwise follow-ups (Tukey's least significant difference) investigating the main effect of model collapsed across the three

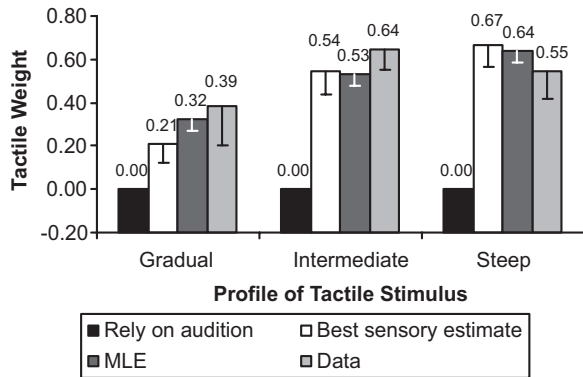


Figure 5. Comparison of weights given to vibrotactile stimuli in bimodal conditions and model predictions for three models of bisensory integration. Mean vibrotactile weights across participants were determined by individual fits to each participant's data, and mean model predictions across participants are based on unimodal judgment uncertainties. Error bars denote standard errors. MLE = maximum likelihood estimation.

levels of tactile profile showed a significant difference between the MLE model and the rely on audition model ( $p = .004$ ) and a marginally significant difference between the best sensory estimate model and the rely on audition model ( $p = .051$ ) but no significant difference between the MLE model and the best sensory estimate model ( $p = .538$ ). The two analyses of tactile weights therefore yielded similar findings: Participants were clearly making use of the tactile stimulus on some or all trials, but the precise manner in which the auditory and tactile stimuli were used could not be determined from the weight data alone.

## Discussion

Our participants judged the temporal order of auditory and vibrotactile stimuli under both unimodal and bimodal conditions. In bimodal conditions, we assessed the influence of the vibrotactile stimulus on the combined judgment by introducing a small discrepancy between the auditory and vibrotactile components of the right-hand combined stimulus. Participants clearly took account of the vibrotactile stimulus when they judged temporal order. We used performance under unimodal conditions to make predictions about bimodal performance based on three models. For weight data, we were able to reject statistically a model in which participants relied entirely on audition to perform the bimodal task but were unable to reject either a strategy that selects the more precise input on each trial or the MLE model, which implies weighted summation of inputs according to their precision. However, when judgment uncertainty was determined in bimodal conditions, we found that participants consistently performed at a higher level than that achieved in either unimodal condition. Only the MLE model predicts this improved level of judgment uncertainty. Furthermore, we were able to reject statistically both alternative models as explanations of our data.

The findings related to observed bimodal judgment uncertainties are particularly important for the following reason. It is possible to mimic the result predicted by the MLE model for sensory weights by the alternative strategy of using only one modality on each trial but using each modality on a proportion of trials determined by the

precision of that modality in unimodal conditions (Ernst & Bühlhoff, 2004). However, this strategy does not predict the increase in precision that is the true hallmark of the MLE model. Thus far, no alternative model has been presented that can predict the low levels of judgment uncertainty observed here, so the MLE model is strongly favored by our data.

Our study represents a formal test of the MLE model of multi-sensory integration for judgments of temporal order. To our knowledge, there have been no studies published previously on this specific issue. The MLE model has been shown to account well for the integration of visual and haptic spatial information (Ernst & Banks, 2002) and for the integration of visual and auditory spatial information (Alais & Burr, 2004). Our data conform with findings from previous studies suggesting that audition dominates over vision (Morein-Zamir et al., 2003; Recanzone, 2003; Scheier et al., 1999; Shams et al., 2000; Welch et al., 1986) and touch (Bresciani et al., 2005) for tasks with a temporal component: Because audition would be expected to have high precision in these tasks, the MLE model can account for its apparent domination, just as it accounts for our audiotactile data. In this study, we have provided a very detailed quantitative analysis of bisensory integration when the relative precision of each modality is varied and have shown that touch can be important under the appropriate circumstances (i.e., when it forms the more reliable input). This finding is consistent with findings of an early study that showed a reciprocal influence of distracting auditory and tactile stimuli on TOJs made in the other modality (Sherrick, 1976).

It may be objected that the rely on audition model is something of a straw man, as very few previous studies have suggested complete auditory dominance over touch for temporal judgments. However, our second alternative model, the best sensory estimate model, is certainly a realistic contender. Like the MLE model, it implies sophisticated knowledge about the precision of each sensory input. Nonetheless, we were able to reject this model and thus favor the MLE model, which suggests that appropriate weighting and combination of information occurs on every trial.

Although ours is the first study to demonstrate statistically optimal integration of bisensory inputs for judgments of temporal order, it is not the first to suggest that TOJs may take into account multiple sources of information in a mathematically sophisticated

Table 2  
Mean Squared Deviation (MSD) of Bimodal Tactile Weights  
From the Predictions of the Models

Model	Gradual tactile profile		Intermediate tactile profile		Steep tactile profile	
	MSD	SE	MSD	SE	MSD	SE
Rely on audition	0.95	0.33	0.62	0.13	0.67	0.16
Best sensory estimate	0.54	0.12	0.26	0.06	0.49	0.13
MLE	0.64	0.20	0.19	0.06	0.36	0.10

Note. Data are shown separately for the combination of the auditory stimulus with each of the three vibrotactile stimuli. MLE = maximum likelihood estimation.

way. For example, prior probability distributions based on previous experience appear to be weighted and integrated with current sensory estimates to perform TOJs (and also anticipate stimulus arrival times) in a Bayesian manner (Miyazaki, Nozaki, & Nakajima, 2005; Miyazaki, Yamamoto, Uchida, & Kitazawa, 2006). In simple terms, participants expect asynchronies that they have repeatedly experienced (although they may also show additional and alternative recalibration effects; see, e.g., Fujisaki, Shimojo, Kashino, & Nishida, 2004, for audiovisual stimuli; Navarra, Soto-Faraco, & Spence, 2007, for audiotactile ones; and Hanson, Heron, & Whitaker, 2008, for all bimodal combinations of vision, audition, and touch). Expectations based on prior experience are easily incorporated into an MLE model of multisensory integration under a single Bayesian framework, so it would be interesting to test for such effects in unison (Ernst & Bühlhoff, 2004).

Because our interpretation relies heavily on changes in the precision with which participants performed the TOJ task, it is important to emphasize that practice effects could not have generated our findings. The order in which participants performed unimodal and bimodal conditions were counterbalanced, so that any improvement (or decrement) over time did not apply to just one phase of the experiment. It might also be objected that we rejected 4 participants prior to obtaining our final sample. These participants were not uniform in terms of the model they best supported, and the motivation for excluding them was the variability they introduced, rather than their conformity to any particular prediction. However, it should be noted that 1 of these 4 participants was an extreme outlier, in that his performance was competent in the unimodal conditions but collapsed almost entirely in the bimodal conditions (which were performed subsequently; i.e., with additional practice). None of our models predicted this pattern, and we wonder whether there was something unusual about this participant that meant an additional redundant stimulus interfered strongly with the first stimulus so it was neither adequately ignored nor usefully integrated.

How might our participants have determined the reliability of each input in order to integrate the inputs appropriately? Participants could not tell in advance of a trial which kind of tactile stimulus they would receive. However, they received only four kinds of stimulus during the experiment (one auditory and three vibrotactile), so it is possible that they were able to build up an accurate estimate of the precision of each kind of stimulus over multiple trials and then classify the tactile component of the bimodal stimulus in order to achieve optimal integration. Given, however, that they received no feedback during the experiment, it is unclear how precision could be accurately determined in this manner. We therefore favor the alternative idea that participants were able to flexibly and near instantaneously estimate the precision of each input on each and every trial (i.e., that the noise of a sensory estimate is represented alongside that estimate in a trial-by-trial manner). This conclusion has been favored by other researchers following demonstrations of MLE integration in which stimuli varied much more widely than was the case here (Hillis et al., 2004).

Turning to possible neurocognitive models for our bimodal TOJ task, we find it tempting to infer from these data that tactile and auditory estimates about time of arrival were integrated for each combined stimulus and that these optimally combined stimuli underwent subsequent comparison to determine their temporal order. This interpretation is intriguing, because it is not immediately obvious how such a process could be accomplished in real time. Sensory inputs are

represented in the brain following a delay that reflects their processing and transmission times. Simple bottom-up models of the TOJ task suggest that separate inputs arrive at some decision center, where they are compared with a more-or-less sophisticated decision rule to determine temporal order (Sternberg & Knoll, 1973; Ulrich, 1987). We can also consider how MLE integration might be accomplished in the brain. One scheme for achieving MLE integration suggests a point-by-point multiplication of two population representations, in which each node corresponds to a particular value and the degree of activation indicates the strength of evidence for that value (Knill & Pouget, 2004). Noisier inputs yield more distributed population responses. For a TOJ task, the values that are represented by different nodes would have to be estimates of time of arrival, and this fact implies that time is no longer represented as time (i.e., in the timing of neural activity) but rather has been converted into a spatial code. Combined sensory estimates would then be compared to evaluate temporal order. This analysis might lead us to reject a real-time account of TOJ task performance and favor the involvement of higher level, potentially post hoc interpretative processes (Dennett & Kinsbourne, 1992).

This, however, is not the only plausible account of our data. It is equally possible that MLE integration occurred at a later stage of processing. In this account, a TOJ is made within each sensory modality first. The left-hand auditory stimulus is compared with the right-hand auditory stimulus, and a parallel comparison occurs for the two tactile stimuli. These comparisons might occur in real time. All that is required is that the output of these comparisons carries quantitative information about the relative timing of left- and right-hand stimuli (i.e., left precedes right by 30 ms, rather than just left precedes right) and that information about precision is produced also. It would then be possible to perform MLE integration in a subsequent computation before reaching a decision.

Previous research assessing MLE models of audiovisual integration for a counting task (Andersen, Tiippana, & Sams, 2004, 2005) has indicated that a model based on continuous representations of sensory inputs (like that assessed here) is superior to a model based on discrete representations (often considered to be a kind of late integration). However, the late integration account outlined above still operates on continuous representations, because the magnitude of the left–right temporal asynchrony is represented following each unimodal comparison. Hence we cannot rule it out with an approach similar to that of Andersen et al. (2005). Although the late integration account does not fit the phenomenology of the task (we were careful to produce stimuli that we felt combined plausibly into bimodal wholes), it remains viable. The interesting issue of whether TOJs are accomplished in real time (i.e., by comparing time of arrival at a decision center) is therefore not resolved by the current data, but it might inform future research comparing multisensory stimuli.

In summary, we have shown that performance on a bimodal TOJ task with combined auditory–tactile stimuli is statistically indistinguishable from the predictions of an optimal MLE model of bimodal integration. Other simple models, which suggested that participants relied exclusively on audition or selected the best unimodal input on each trial, were rejected on the basis of reliable differences between model predictions and observed behavior. We therefore conclude that although the locus of integration remains uncertain, humans nonetheless are able to integrate auditory and vibrotactile information in a statistically optimal manner while determining the temporal order of bisensory events.

## References

- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14*, 257–262.
- Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Brain Research: Cognitive Brain Research, 21*, 301–308.
- Andersen, T. S., Tiippana, K., & Sams, M. (2005). Maximum likelihood integration of rapid flashes and beeps. *Neuroscience Letters, 380*, 155–160.
- Bertelson, P., & de Gelder, B. E. M. (2004). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and cross-modal attention* (pp. 141–178). Oxford, England: Oxford University Press.
- Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory–visual spatial discordance. *Perception and Psychophysics, 29*, 578–584.
- Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. *Journal of Vision, 6*, 554–564.
- Bresciani, J. P., & Ernst, M. O. (2007). Signal reliability modulates auditory–tactile integration for event counting. *NeuroReport, 18*, 1157–1161.
- Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: Auditory signals can modulate tactile tap perception. *Experimental Brain Research, 162*, 172–180.
- de Lange, H. (1958). Research into the dynamic nature of the human fovea–cortex systems with intermittent and modulated light: I. Attenuation characteristics with white and colored light. *Journal of the Optical Society of America, 48*, 777–784.
- Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences, 15*, 183–247.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*, 429–433.
- Ernst, M. O., & Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*, 162–169.
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*, 773–778.
- Hanson, J. V., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Experimental Brain Research, 185*, 347–352.
- Hawken, M. J., Shapley, R. M., & Grosz, D. H. (1996). Temporal frequency selectivity in monkey visual cortex. *Visual Neuroscience, 13*, 477–492.
- Heron, J., Whitaker, D., & McGraw, P. V. (2004). Sensory uncertainty governs the extent of audio–visual interaction. *Vision Research, 44*, 2875–2884.
- Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision, 4*, 967–992.
- Kitazawa, S., Moizumi, S., Okuzumi, A., Saito, F., Shibuya, S., et al. (2007). Reversal of subjective temporal order due to sensory and motor integrations. In P. Haggard, M. Kawato, & Y. Rossetti (Eds.), *Attention and performance XXII: Sensorimotor foundations of higher cognition* (pp. 73–97). Oxford, England: Oxford University Press.
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27*, 712–719.
- Miyazaki, M., Nozaki, D., & Nakajima, Y. (2005). Testing Bayesian models of human coincidence timing. *Journal of Neurophysiology, 94*, 395–399.
- Miyazaki, M., Yamamoto, S., Uchida, S., & Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nature Neuroscience, 9*, 875–877.
- Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Brain Research: Cognitive Brain Research, 17*, 154–163.
- Müller, J. (1838). *Handbook der Physiologie des Menschen* [Handbook of human physiology] (Vol. 2). Koblenz, Germany: J. Hölscher.
- Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neuroscience Letters, 413*, 72–76.
- Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception and Psychophysics, 6*, 203–205.
- Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology, 89*, 1078–1093.
- Roach, N. W., Heron, J., & McGraw, P. V. (2006). Resolving multisensory conflict: A strategy for balancing the costs and benefits of audio–visual integration. *Proceedings of the Royal Society of London, Series B: Biological Sciences, 273*, 2159–2168.
- Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science, 143*, 594–596.
- Rosenberger, W. F., & Grill, S. E. (1997). A sequential design for psychophysical experiments: An application to estimating timing of sensory events. *Statistics in Medicine, 16*, 2245–2260.
- Scheier, C. R., Nijhawan, R., & Shimojo, S. (1999). Sound alters visual temporal resolution. *Investigative Ophthalmology and Visual Science, 40*, 169.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature, 408*, 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Brain Research: Cognitive Brain Research, 14*, 147–152.
- Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport, 16*, 1923–1927.
- Sherrick, C. E. (1976). The antagonisms of hearing and touch. In S. K. Hirsh, D. H. Eldredge, I. J. Hirsh, & S. R. Silverman (Eds.), *Hearing and Davis: Essays honoring Hallowell Davis* (pp. 149–158). Saint Louis, MI: Washington University Press.
- Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130*, 799–832.
- Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 629–686). London: Academic Press.
- Ulrich, R. (1987). Threshold models of temporal-order judgments evaluated by a ternary response task. *Perception and Psychophysics, 42*, 224–239.
- Van Beers, R. J., Sittig, A. C., & Gon, J. J. (1999). Integration of proprioceptive and visual position information: An experimentally supported model. *Journal of Neurophysiology, 81*, 1355–1364.
- Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *Journal of the Acoustical Society of America, 66*, 1364–1380.
- Wada, Y., Kitagawa, N., & Noguchi, K. (2003). Audio–visual integration in temporal perception. *International Journal of Psychophysiology, 50*, 117–124.
- Weisenberger, J. M. (1986). Sensitivity to amplitude-modulated vibrotactile signals. *Journal of the Acoustical Society of America, 80*, 1707–1715.
- Welch, R. B., DuttonHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception and Psychophysics, 39*, 294–300.
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*, 638–667.
- Welch, R. B., Widawski, M. H., Harrington, J., & Warren, D. H. (1979). An examination of the relationship between visual capture and prism adaptation. *Perception and Psychophysics, 25*, 126–132.
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics, 63*, 1293–1313.
- Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of Vision, 8*(3), 1–11.

(Appendix follows)

Appendix

Predictions for Individual Participants for Each of the Three Models, Alongside Data Obtained in Bimodal Conditions

Table A1  
*Predicted Tactile Weights Derived From the Rely on Audition and Best Sensory Estimate Models*

Subject <sup>a</sup>	Age	Order <sup>b</sup>	Rely on audition model		Best sensory estimate model					
			All tactile profiles		Gradual tactile profile		Intermediate tactile profile		Steep tactile profile	
			Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval
1	20	1alt2	0	0, 0	0	0, 1	0	0, 1	0	0, 1
2	21	21alt	0	0, 0	0	0, 1	0	0, 1	0	0, 1
3	27	1t1a2	0	0, 0	0	0, 0	0	0, 1	1	0, 1
4	24	21alt	0	0, 0	0	0, 1	1	0, 1	0	0, 1
5	29	1alt2	0	0, 0	0	0, 1	1	0, 1	1	0, 1
6	24	1t1a2	0	0, 0	0	0, 0	0	0, 0	0	0, 1
7	23	21t1a	0	0, 0	0	0, 0	0	0, 0	0	0, 0
8	25	1alt2	0	0, 0	0	0, 1	1	0, 1	1	1, 1
9	29	1t1a2	0	0, 0	0	0, 1	0	0, 1	1	0, 1
10	22	21t1a	0	0, 0	0	0, 0	0	0, 1	0	0, 1
11	28	21t1a	0	0, 0	0	0, 0	0	0, 1	0	0, 1
12	23	1t1a2	0	0, 0	0	0, 1	1	0, 1	1	0, 1
13	22	21alt	0	0, 0	1	1, 1	1	0, 1	1	1, 1
14	20	21t1a	0	0, 0	0	0, 1	0	0, 1	0	0, 1
15	25	1alt2	0	0, 0	0	0, 1	1	0, 1	1	0, 1
16	33	21alt	0	0, 0	0	0, 1	1	1, 1	1	0, 1
17	30	1t1a2	0	0, 0	0	0, 1	1	0, 1	1	0, 1
18	24	21t1a	0	0, 0	1	0, 1	1	0, 1	1	1, 1
19	32	21alt	0	0, 0	1	0, 1	1	0, 1	1	0, 1
20	26	21t1a	0	0, 0	1	1, 1	1	0, 1	1	1, 1
21	20	1t1a2	0	0, 0	0	0, 1	1	0, 1	1	1, 1
22	28	21alt	0	0, 0	1	0, 1	1	1, 1	1	1, 1
23	30	1alt2	0	0, 0	0	0, 0	0	0, 1	1	0, 1
24	29	1alt2	0	0, 0	0	0, 0	0	0, 1	1	1, 1
25	22	21alt	0	0, 0	1	0, 1	1	1, 1	1	0, 1
26	29	1alt2	0	0, 0	0	0, 0	0	0, 1	1	0, 1
27	23	1t1a2	0	0, 0	0	0, 1	1	1, 1	1	1, 1
28	19	1alt2	0	0, 0	0	0, 0	0	0, 0	0	0, 0

Note. Confidence intervals are based on a percentile parametric bootstrap method using 4,999 simulations.  
<sup>a</sup> Only participants 1–24 were included in the main analysis. <sup>b</sup> For experimental order, 1a = unimodal auditory, 1t = unimodal tactile, 2 = bimodal.

Table A2

*Predicted Tactile Weights Derived From the MLE Model, Alongside Observed Bimodal Data*

Subject <sup>a</sup>	MLE model						Bimodal data					
	Gradual tactile profile		Intermediate tactile profile		Steep tactile profile		Gradual tactile profile		Intermediate tactile profile		Steep tactile profile	
	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval
1	0.47	0.12, 0.85	0.25	0.05, 0.66	0.28	0.07, 0.67	0.46	-0.45, 1.33	0.98	0.44, 1.54	0.44	-0.38, 1.07
2	0.10	0.00, 0.53	0.25	0.02, 0.69	0.20	0.02, 0.62	0.64	-1.10, 3.92	-0.28	-5.24, 2.39	-0.04	-3.60, 12.25
3	0.08	0.01, 0.29	0.23	0.04, 0.55	0.60	0.21, 0.90	0.32	-0.60, 0.98	-0.18	-1.21, 0.62	0.22	-0.61, 0.84
4	0.25	0.04, 0.58	0.62	0.22, 0.89	0.33	0.08, 0.68	0.58	-0.13, 1.28	0.40	-0.25, 1.22	1.06	0.13, 1.89
5	0.22	0.03, 0.70	0.73	0.39, 0.96	0.80	0.43, 0.97	0.88	-0.01, 2.40	0.88	-0.19, 1.92	-0.64	-1.38, 0.35
6	0.08	0.00, 0.30	0.16	0.04, 0.28	0.48	0.12, 0.79	0.70	-0.01, 1.38	0.84	0.26, 1.36	1.00	0.49, 1.41
7	0.01	0.00, 0.07	0.02	0.00, 0.12	0.11	0.01, 0.38	-0.64	-1.59, 0.35	0.58	-0.18, 1.34	-0.60	-1.00, -0.21
8	0.49	0.18, 0.86	0.62	0.26, 0.92	0.90	0.71, 0.99	-0.06	-0.71, 0.75	0.36	0.16, 0.89	0.70	0.14, 1.27
9	0.03	0.00, 0.56	0.41	0.02, 0.90	0.81	0.44, 0.98	-0.42	-2.45, 1.11	0.88	-0.88, 3.18	-0.14	-1.17, 1.05
10	0.19	0.02, 0.47	0.27	0.05, 0.62	0.34	0.06, 0.67	0.90	0.00, 1.66	0.04	-0.78, 0.93	0.82	0.15, 1.39
11	0.07	0.00, 0.25	0.44	0.04, 0.90	0.27	0.02, 0.69	-0.24	-0.66, 0.25	0.26	-0.13, 0.68	0.58	0.25, 0.93
12	0.49	0.15, 0.90	0.82	0.45, 0.98	0.77	0.43, 0.97	0.50	-0.27, 1.26	0.80	-0.34, 1.51	0.98	0.19, 1.87
13	0.86	0.53, 0.98	0.81	0.47, 0.98	0.93	0.70, 0.99	1.40	0.31, 2.23	0.92	0.22, 1.63	0.88	0.11, 1.82
14	0.20	0.03, 0.51	0.39	0.08, 0.76	0.41	0.11, 0.78	-0.20	-0.88, 0.40	0.52	-0.05, 0.89	-0.34	-0.98, 0.36
15	0.30	0.00, 0.96	0.52	0.11, 0.98	0.77	0.30, 0.99	-0.54	-2.48, 1.24	1.62	0.22, 3.19	0.54	-0.10, 0.97
16	0.22	0.00, 1.00	0.91	0.61, 1.00	0.88	0.42, 1.00	-1.22	-12.8, 3.95	1.40	0.15, 4.02	1.78	0.86, 2.89
17	0.28	0.03, 0.73	0.74	0.36, 0.97	0.77	0.40, 0.97	-0.38	-1.35, 0.86	1.02	0.69, 1.41	0.50	-0.01, 1.10
18	0.50	0.18, 0.89	0.73	0.38, 0.96	0.85	0.57, 0.98	2.30	0.32, 5.10	0.98	0.11, 1.63	1.54	0.71, 2.40
19	0.62	0.14, 0.92	0.83	0.49, 0.98	0.73	0.27, 0.95	2.56	1.11, 3.78	0.90	-0.34, 1.68	1.30	0.37, 2.08
20	0.93	0.72, 1.00	0.69	0.25, 0.96	0.82	0.50, 0.98	0.72	-0.20, 1.44	0.84	0.09, 1.56	0.56	-0.23, 1.37
21	0.43	0.05, 0.96	0.79	0.36, 0.99	0.93	0.68, 1.00	0.76	-0.80, 2.54	0.06	-0.93, 1.19	-0.12	-2.68, 1.80
22	0.83	0.35, 1.00	0.88	0.53, 1.00	0.96	0.83, 1.00	0.38	-0.37, 1.10	0.58	0.07, 1.23	0.56	0.03, 1.36
23	0.08	0.01, 0.30	0.29	0.04, 0.64	0.56	0.20, 0.86	0.86	-0.10, 1.71	0.54	-0.47, 1.50	0.70	-0.11, 1.44
24	0.03	0.00, 0.38	0.39	0.05, 0.80	0.83	0.52, 0.97	-0.98	-6.19, 2.63	0.50	-0.92, 1.84	0.82	-0.20, 2.46
25	0.90	0.28, 1.00	0.96	0.57, 1.00	0.95	0.48, 1.00	-0.76	-17.1, 3.33	1.32	-0.58, 2.48	0.36	-0.55, 1.76
26	0.02	0.00, 0.46	0.08	0.00, 0.65	0.78	0.39, 0.98	-0.80	-2.38, 3.52	0.04	-1.67, 1.82	-0.02	-1.14, 1.23
27	0.31	0.00, 0.97	0.92	0.63, 1.00	0.92	0.58, 1.00	-8.04	-2.623, 53.40	1.38	-301, 121.11	1.62	-6.16, 22.21
28	0.00	0.00, 0.01	0.02	0.01, 0.09	0.03	0.01, 0.11	0.92	-0.16, 1.62	0.38	-1.26, 1.62	-0.08	-1.80, 1.48

*Note.* Confidence intervals are based on a percentile parametric bootstrap method using 4,999 simulations.

<sup>a</sup> Only participants 1–24 were included in the main analysis.

(Appendixes continue)

Table A3

*Predicted Judgment Uncertainties Derived From the Rely on Audition and Best Sensory Estimate Models*

Subject <sup>a</sup>	Age	Order <sup>b</sup>	Rely on audition model		Best sensory estimate model					
			All tactile profiles		Gradual tactile profile		Intermediate tactile profile		Steep tactile profile	
			Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval
1	20	1a1t2	118	58, 196	118	50, 158	118	58, 181	118	58, 178
2	21	21a1t	154	78, 296	154	78, 287	154	78, 266	154	78, 272
3	27	1t1a2	79	35, 108	79	35, 108	79	35, 105	65	24, 82
4	24	21a1t	87	40, 124	87	40, 120	68	28, 91	87	39, 115
5	29	1a1t2	143	72, 290	143	72, 251	88	38, 121	72	35, 110
6	24	1t1a2	94	43, 134	94	43, 134	94	43, 134	94	39, 111
7	23	21t1a	37	14, 57	37	14, 57	37	14, 57	37	14, 56
8	25	1a1t2	113	56, 194	113	49, 136	88	37, 117	37	9, 51
9	29	1t1a2	254	125, 695	254	125, 647	254	117, 418	122	62, 198
10	22	21t1a	64	24, 86	64	24, 86	64	24, 84	64	24, 81
11	28	21t1a	33	7, 45	33	7, 45	33	4, 40	33	6, 44
12	23	1t1a2	163	82, 389	163	72, 218	77	36, 123	89	42, 129
13	22	21a1t	110	56, 200	45	16, 73	53	17, 79	30	10, 51
14	20	21t1a	72	32, 108	72	31, 105	72	30, 97	72	29, 91
15	25	1a1t2	147	68, 868	147	68, 317	140	65, 215	81	43, 132
16	33	21a1t	493	198, 21,056	493	198, 2,178	153	153, 167	183	103, 332
17	30	1t1a2	130	63, 253	130	63, 212	78	31, 108	71	29, 100
18	24	21t1a	168	89, 346	167	77, 216	102	49, 155	71	33, 107
19	32	21a1t	228	116, 469	178	88, 271	104	46, 158	138	74, 228
20	26	21t1a	202	99, 523	54	22, 85	136	66, 211	94	41, 135
21	20	1t1a2	269	130, 1,154	269	124, 457	139	70, 228	76	33, 121
22	28	21a1t	330	152, 1,818	149	73, 287	119	59, 213	64	28, 96
23	30	1a1t2	84	37, 118	84	37, 118	84	37, 113	75	30, 88
24	29	1a1t2	187	95, 349	187	94, 348	187	91, 272	84	40, 127
25	22	21a1t	882	240, 495,620	286	138, 618	176	83, 343	202	102, 428
26	29	1a1t2	233	121, 684	233	121, 650	233	121, 565	125	62, 202
27	23	1t1a2	417	182, 2,268	417	166, 1,058	120	61, 200	127	67, 225
28	19	1a1t2	19	19, 20	19	19, 20	19	19, 20	19	19, 20

*Note.* Confidence intervals are based on a percentile parametric bootstrap method using 4,999 simulations.

<sup>a</sup> Only participants 1–24 were included in the main analysis. <sup>b</sup> For experimental order; 1a = unimodal auditory; 1t = unimodal tactile, 2 = bimodal.

Table A4

*Predicted Judgment Uncertainties Derived From the MLE Mode, Alongside Observed Bimodal Data*

Subject <sup>a</sup>	MLE model						Bimodal data					
	Gradual tactile profile		Intermediate tactile profile		Steep tactile profile		Gradual tactile profile		Intermediate tactile profile		Steep tactile profile	
	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval	Estimate	95% confidence interval
1	86	44, 122	102	54, 143	100	53, 139	80	50, 96	58	36, 76	77	49, 95
2	146	76, 256	133	73, 220	138	75, 219	211	143, 358	258	170, 562	283	167, 4620
3	76	35, 99	69	33, 87	50	22, 63	70	46, 90	81	51, 99	51	33, 68
4	75	38, 99	54	26, 69	71	36, 90	86	60, 116	65	42, 80	82	51, 102
5	126	68, 203	75	36, 99	64	33, 91	120	92, 164	97	61, 114	74	49, 94
6	90	42, 124	86	42, 114	68	35, 86	62	38, 71	48	29, 56	39	21, 46
7	37	14, 56	37	14, 55	35	14, 50	98	67, 132	66	39, 77	43	29, 56
8	81	43, 105	69	34, 92	35	9, 46	56	35, 74	31	14, 37	35	19, 46
9	250	123, 590	194	103, 353	110	59, 163	205	129, 648	177	115, 275	108	73, 158
10	58	23, 74	55	23, 66	52	23, 64	76	47, 93	58	38, 76	53	33, 65
11	32	7, 41	25	4, 31	28	6, 35	36	21, 43	28	15, 35	28	16, 35
12	116	64, 171	70	34, 102	78	40, 107	72	45, 84	75	49, 96	75	48, 91
13	42	15, 61	48	17, 66	29	10, 45	92	62, 122	79	51, 103	70	45, 90
14	64	30, 87	56	28, 75	55	27, 69	60	37, 71	59	40, 73	53	33, 68
15	123	63, 273	101	56, 179	71	40, 112	169	110, 710	93	59, 199	66	44, 94
16	435	187, 1946	146	123, 166	172	98, 294	354	214, 8,397	154	102, 241	114	74, 171
17	110	59, 172	67	29, 88	62	29, 82	118	85, 180	44	37, 52	42	24, 49
18	118	67, 168	87	46, 126	65	32, 94	135	90, 231	60	37, 74	81	50, 94
19	140	79, 219	95	44, 131	118	68, 180	152	102, 223	89	59, 117	76	46, 94
20	52	22, 78	113	60, 173	85	40, 116	81	54, 109	54	33, 68	62	42, 77
21	204	111, 377	123	66, 192	73	32, 111	140	93, 213	114	77, 153	141	95, 272
22	136	70, 240	112	57, 183	63	28, 91	63	38, 74	44	26, 55	49	30, 65
23	81	37, 109	71	35, 92	56	27, 67	82	54, 109	89	54, 109	67	42, 79
24	184	93, 326	146	82, 220	77	39, 107	215	145, 1,109	141	94, 208	113	83, 153
25	272	132, 553	173	81, 317	197	99, 390	236	142, 3,146	175	122, 283	114	77, 149
26	231	119, 592	223	119, 517	110	59, 168	154	97, 253	159	107, 232	103	67, 135
27	346	149, 944	115	61, 180	121	65, 201	668	340, 603,440	522	290, 289,290	323	196, 13,317
28	19	19, 20	19	18, 20	19	18, 20	135	104, 186	153	115, 256	136	91, 193

*Note.* Confidence intervals are based on a percentile parametric bootstrap method using 4,999 simulations.

<sup>a</sup> Only participants 1–24 were included in the main analysis.

Received November 26, 2007  
Revision received September 23, 2008  
Accepted October 13, 2008 ■