



Perceptual uncertainty and line-call challenges in professional tennis

George Mather*

Department of Psychology, University of Sussex, Falmer, Brighton BN1 9QH, UK

Fast-moving sports such as tennis require both players and match officials to make rapid accurate perceptual decisions about dynamic events in the visual world. Disagreements arise regularly, leading to disputes about decisions such as line calls. A number of factors must contribute to these disputes, including lapses in concentration, bias and gamesmanship. Fundamental uncertainty or variability in the sensory information supporting decisions must also play a role. Modern technological innovations now provide detailed and accurate physical information that can be compared against the decisions of players and officials. The present paper uses this psychophysical data to assess the significance of perceptual limitations as a contributor to real-world decisions in professional tennis. A detailed analysis is presented of a large body of data on line-call challenges in professional tennis tournaments over the last 2 years. Results reveal that the vast majority of challenges can be explained in a direct highly predictable manner by a simple model of uncertainty in perceptual information processing. Both players and line judges are more reliable than players. Judgements are more difficult for balls bouncing near base and service lines than those bouncing near side and centre lines. There is no evidence for significant errors in localization due to image motion.

Keywords: psychophysics; visual perception; perceptual judgement

1. INTRODUCTION

Everyday experience indicates that human perceptual judgement is so accurate and reliable that disagreements between observers are rare. Sport offers a prime example: players can compete (and officials can adjudicate) only because they are able to make fast accurate perceptual judgements about the position, speed and direction of relevant objects in the visual scene such as their opponent or the ball in play. However, disputes can and do arise between players and officials, typically involving judgements about whether a ball bounced on or crossed a line on the field or court. Several factors must contribute to these disagreements: the official may suffer a lapse in concentration; a player may have a strong disposition in favour of certain decisions or may be engaging in gamesmanship; or the disagreement may originate in the participants' perceptual information processing systems. A great deal of laboratory research has been conducted on the perceptual capacities of human observers. A consistent finding in this research is that, while humans can make perceptual judgements with extreme precision, performance ultimately hits limits set by the neural machinery serving vision. For example, neurons at all levels in the visual system from retina to visual cortex can be considered to take discrete spatial samples of the visual image, introducing an unavoidable degree of uncertainty regarding the precise position of spatial features in the image (see Klein & Levi 1987; Watt & Hess 1987; Wilson 1991; Land & Nilsson 2002). Furthermore, neural processing is prone to internal noise, or random fluctuations in response (see Barlow 1956; Pelli 1991; White et al. 2000; Li et al. 2006). Consequently, repeated presentation

of the same physical stimulus will never produce an identical neural response. If these sensory limits have a significant impact on perception, we should expect to find evidence for them in the performance of demanding realworld tasks such as judgements in professional sport. This paper reports an investigation of perceptual uncertainty in line calls in professional tennis.

Professional tennis tournaments organized by the Association of Tennis Professionals (ATP) use the Hawk-Eye ball-tracking system that can locate the threedimensional position of the ball in play to within 3 mm (see Fischetti 2007). Multiple video cameras are trained on the court, and image processing software computes the three-dimensional location of the ball in each video frame in order to recover its trajectory. Line calls are still made visually by line judges who decide whether the ball played in each stroke bounced inside or outside the court. Players in ATP tournament matches may challenge line calls (no more than two unsuccessful challenges per set). When a challenge is made, the umpire calls for a review of the Hawk-Eye data to determine whether the ball actually did bounce on the inside or outside of the court line. The umpire either upholds or overturns the line judge's call accordingly. The details of each challenge are recorded on a pro forma by match officials. Actual ball bounce position relative to the court line, as assigned by Hawk-Eye, is recorded to the nearest millimetre along with the judge's decision and other details of the call.

Challenge records can be treated as psychophysical data, since they specify the relation between a physical event (ball bounce location) and a perceptual judgement ('in' or 'out' call), and can be analysed to investigate whether challenges generally reflect gamesmanship, lapses in concentration or genuine perceptual uncertainty. Can



Figure 1. Line-call challenges and errors in recent ATP tennis tournaments. Filled circles, the proportion of challenges to line judge calls occurring at each ball bounce position. Negative positions indicate balls bouncing inside the court line, and positive positions indicate those bouncing outside the court line. Open circles, the proportion of line judge errors in challenged calls at each ball position. Data were taken from ATP records and grouped into 5 mm position bins for analysis.

line-call challenge and error data be explained using the concepts developed in laboratory studies of visual perceptual limits?

2. DATASET

In response to a request, the ATP provided the author with records of all 1473 challenges made by 246 professional tennis players (or doubles teams) during 15 ATP tennis tournaments worldwide during 2006 and 2007. The following ATP events were included in the dataset: Beijing (2006), Cincinnati (2006), Indian Wells (2007), Indianapolis (2006), Los Angeles (2006), Madrid (2006), Miami (2006, 2007), Moscow (2006), New Haven (2006), Paris (2006), Rotterdam (2007), Shanghai (2006), Toronto (2006) and Washington (2006). The details of each challenge were entered into a spreadsheet and analysed using VISUAL BASIC macros and MATLAB scripts. Hawk-Eye ball position data were grouped into 5 mm bins for the purposes of analysis.

Challenges predominantly occurred when the ball bounced close to a court line: 94% of the challenges occurred for balls bouncing within 100 mm of the line, a distance less than twice the diameter of the ball. The remaining 6% of the challenges were spread thinly over a wide range of ball positions greater than 100 mm; the largest distance involved in a challenge was 449 mm (in this instance, the line judge's call was good). Challenges at large distances may arise from lapses in judgement or gamesmanship, but their wide dispersion and very low number preclude meaningful statistical analysis. The analysis presented here will focus on the overwhelming majority of calls (1380) involving balls within 100 mm of a court line.

The filled circles in figure 1 show the proportion of these challenges falling in each 5 mm position bin (negative values indicate balls bouncing inside the court line, and positive values indicate balls bouncing outside the court line). Challenges rise rapidly and nonlinearly as the bounce of the ball approaches the line from either side, with a peak for balls bouncing within 5 mm of the line. Challenged decisions were equally split between out and in calls by the line judge (51.3% and 48.7%, respectively), indicating that the greater distance of the player from the ball in out calls does not lead to more challenges. Over 60% of calls were correct, indicating that judgements were based on relevant visual information. But a significant proportion (39.3%) of the challenged line calls turned out to be incorrect according to the Hawk-Eye data. The open circles in figure 1 plot the proportion of these line judge errors that fell in each 5 mm position bin. Errors also rise nonlinearly as the ball approaches the line from either side.

The sharp peak in challenges and errors at the court line suggests that they involve balls that are 'too close to call', rather than lapses in concentration or gamesmanship. A simple psychophysical model of perceptual uncertainty was developed to assess how well the variation in challenges and errors can be explained by known limitations in perceptual processing.

3. MODEL

The model is illustrated schematically in figure 2a. Following a ball bounce (box 1), the player and the line judge each compute the ball bounce position relative to the court line (boxes 2 and 3), based on the information furnished by the visual system. If the sign of the player's assignment (IN versus OUT) agrees with the official's assignment as indicated by his/her call (box 4), then the sequence ends (9). If the player and the line judge assignments disagree, the player challenges the call (box 5). If Hawk-Eye data indicates that the player's assignment is correct (box 6), then the challenge is successful (box 7) and the line judge's call is recorded as an error. If the player's assignment is incorrect, then the challenge fails (box 8).

The critical aspect of the model is the mechanism by which the player and the line judge each compute ball bounce position relative to the line (boxes 2 and 3), based on visual information. A core assumption in modern sensory science, as mentioned in §1, is that neural signals are subject to internal 'noise', or random fluctuations in output. For example, psychophysical laboratory studies of visual acuity reveal that human observers are extremely good at judging the relative position of elements in the visual image (Westheimer 1981) but performance is limited by a degree of intrinsic uncertainty regarding position, which has been attributed to the sampling properties of the retina in the eye and to neural noise perturbing position information during processing in the brain (Klein & Levi 1987; Watt & Hess 1987; Wilson 1991). On this basis, we can assume that player and line judge position assignments are perturbed by intrinsic perceptual uncertainty, so that a certain proportion of their assignments will be incorrect.

Although the situation is undoubtedly complex, with multiple sources of error affecting the decisions of both players and officials, the degree of order evident in the data plotted in figure 1 hints at a relatively simple account of how positional uncertainty contributes to challenges and errors. The aim here was to develop the simplest possible psychophysical model, and evaluate how well it can account for this outwardly complex real-world



Figure 2. A simple psychophysical model of line-call challenges and errors. (*a*) Flow chart describing the logical flow and structure of the model. Following a ball bounce, (1) the player and judge each compute bounce position relative to the court line (2 and 3); the player compares his/her own assignment with the judge's (4); if they agree, the sequence ends (9) otherwise the player challenges (5); Hawk-Eye data are used to assess whether the challenge is correct (6); and the challenge is upheld (7) or rejected (8). (*b*) Computation of ball bounce position (boxes 2 and 3 in the flow chart), illustrated for a ball bouncing 20 mm outside the line. The visual system assigns positions to the line and the ball, but each assignment is subject to a Gaussian uncertainty distribution. The solid and dashed curves represent the possible assignments for the line and the ball, respectively. For some calls (grey line), the assignments lead to a correct decision, while for others (black line) they lead to an incorrect decision. In this example, the space constant of each distribution is 30 mm (the best-fitting value for the player, according to the modelling results in figure 3).

perceptual decision. It was assumed that each position assignment of the ball bounce and the court line in boxes 2 and 3 is perturbed according to a Gaussian uncertainty distribution centred on the correct position (Gaussian uncertainty distributions were suggested by Wilson (1991) and Watt & Hess (1987)). In the interests of simplicity, we assumed similar uncertainty distributions for ball bounce and line position, though in reality ball bounce position may be subject to some (unknown) greater degree of uncertainty due to the ball's motion. All sources of possible contributory noise were represented in a single uncertainty distribution. We also assumed that there is no substantial bias in either the player's or official's judgements, because the error distributions in figure 1 peak close to a position of zero (see below).

Figure 2b shows a simulation of the player's position assignment (box 2) for ball bounces 20 mm outside the court. The position assigned to the line for any given bounce is drawn randomly from a Gaussian distribution centred on 0 mm (solid curve), and the position assigned to the ball is



Figure 3. Predictions of a simple psychophysical model of challenges and errors. Filled and open circles show challenge and error data, respectively, re-plotted from figure 1. Lines show the predictions of a Monte Carlo simulation (described in detail in the text). (*a*) Predicted challenges and (*b*) predicted errors. Predictions are based on player and judge positional uncertainty space constants of 30 and 22 mm, respectively.

drawn from another distribution centred on +20 mm (dashed curve). The grey lines show one possible outcome: the line and the ball are assigned positions corresponding to -10 and +30 mm, respectively, so that the player correctly decides that the ball was out. The black lines show another possible outcome: the line and the ball are assigned positions corresponding to +15 and +5 mm, respectively, so that the player incorrectly decides that the ball was in. The probability of an incorrect assignment should decline as the ball moves further from the line, for a given uncertainty space constant, because the two distributions move further apart. Corresponding distributions can be drawn for the line judge's position assignments (not shown) leading to his/her decision regarding in versus out. Once the player's and official's position assignments have each been computed, the simple logic in figure 2a determines the consequences.

A Monte Carlo simulation of this simple model was implemented in MATLAB. The sequence illustrated in figure 2a was executed 25 000 times at each of 20 ball bounce positions up to 100 mm on either side of the line, to compute challenge and error probabilities at each position. Different random samples from the uncertainty distributions were drawn in each run. The only two free parameters in the model were the space constants of the player's and judge's positional uncertainty distributions. The whole simulation was repeated for different combinations of player and line judge space constants. Bestfitting combinations of space constants were selected on the basis of values that minimized the root mean square error of predictions against the data in figure 1.

4. RESULTS AND DISCUSSION

(a) Challenge and error rates

The Monte Carlo simulation showed that, for the bestfitting space constants, the condition for a challenge (player and judge disagree) was met in 20.8% of all trials. In 39.6% of the simulated challenges, a line judge error was recorded, which is very close to the 39.3% of errors found in the actual challenge records. The symbols in figure 3 re-plot data from figure 1, collapsed across positive and negative positions because the distributions are symmetrical. The smooth curves in the figure show predictions from the model. Figure 3a represents the proportion of the simulated challenges at each ball bounce position. Figure 3b shows the proportion of simulated



Figure 4. Psychometric function relating the proportion of 'out' calls by the line judge to ball bounce position. Negative positions indicate balls bouncing inside the court line, and positive positions indicate those bouncing outside the court line. The smooth curve was fitted by the maximum-likelihood method. The size of each symbol reflects the number of challenges recorded at each position (and therefore the contribution of each point to the curve fit).

errors at each ball bounce position. For the best-fitting lines, the space constants of the player's and judge's uncertainty distribution were 30 and 22 mm, respectively. The model provides a very accurate account of the data. The coefficient of determination (r^2) is 0.963 between the model and challenge data, and 0.938 between the model and error data. One limitation of the modelling is that it grouped together challenges to in and out calls, but one might expect a wider space constant for the latter since the player is likely to be farther from the ball.

(b) Localization errors

Psychophysical laboratory studies have shown that retinal motion biases perceived position. At any one instant, a moving object may appear displaced in position towards its direction of travel (see Krekelberg & Lappe 2000; Nijhawan 2002; Öğmen *et al.* 2004; Bressler & Whitney 2006; Linares *et al.* 2007). Does this effect contribute to the errors made by line judges? If so, one would expect that balls travelling from the in side of the court towards the out side, and bouncing just inside the line, will tend to be called out. Figure 4 shows a psychometric function



Figure 5. Predictions and data for (*a*) challenges and (*b*) errors for three subsets of the data, representing base/service line bounces (black), side line bounces (light grey) and centre line bounces (dark grey). Best-fitting space constants for these fits are given in table 1.Squares, base/service challenges and errors; circles, side challenges and errors; diamonds, centre challenges and errors, respectively, in (*a*) and (*b*).

from the empirical data relating the proportion of out calls to the position of the ball bounce. On the basis of laboratory studies, the prediction is that the midpoint (0.5) of the best-fitting function should be located at a negative position. A logistic psychometric function was fitted using the *psignifit* toolbox v. 2.5.6 for MATLAB, which implements the maximum-likelihood method described by Wichmann & Hill (2001). The midpoint of the function is located just inside the court at a position of -14.4 mm, consistent with the prediction. However, the effect is very small, amounting to only one-third of the width of the ball's contact patch with the court (44 mm according to Fischetti 2007).

(c) The effect of court location

The data in figure 4 reflect a diverse range of conditions, particularly with regard to the trajectory of the ball relative to the court line, and this may have obscured any effect of movement direction on apparent bounce position. Bounces near some court lines may be more difficult to judge than those near other court lines. Fortunately, challenge records include information specifying the court line involved in each decision, allowing the challenge data to be partitioned into three subsets: base/service line challenges (these lines run sideways across the court, 628 challenges in the dataset); side line challenges (these lines run lengthways down each side of the court, 567 challenges); and centre line challenges (these lines run lengthways across the net, 185 challenges). Separate simulations and curve fits identical to those described above were performed on each of the three data subsets.

Data points in figure 5a show the proportion of challenges and in figure 5b line judge errors in each data subset, as a function of ball bounce position. The solid curves show the predictions of Monte Carlo simulations applied separately to each subset. The average coefficient of determination (r^2) for these six curves was 0.86 (s.d. = 0.12). Best-fitting curves were identical for the centre and side line challenges. Base/service line challenges and errors are spread further from the line. Table 1 summarizes the best-fitting parameters for the simulations and psychometric functions (all values in millimetres).

Line judges are generally more precise at locating ball bounce position than players, as indicated by the smaller estimated space constant for their uncertainty distributions.

Table 1. Parameters of the best-fitting functions for three subsets of the data.

line	player, σ	line judge, σ	P50	threshold
base/service	46	32	-9.15	96.33
side	26	14	-13.37	38.79
centre	26	14	4.33	59.11

However, the variation in space constant indicates that both players and line judges find bounces near the base/service lines much harder to judge than bounces near the side and centre lines. Best-fitting psychometric functions show some small differences in 50% points (P50), but statistical comparisons performed using the *psignifit* toolbox indicate that the differences between the functions are not significant. They offer no support for the idea that ball bounce position is mis-localized by line judges.

Why are bounces near the base and service lines harder to judge than those at the side and centre lines? Line judges sit at a fixed distance from the court, collinear with the line they are calling and much closer to the base and service lines than to the side and centre lines (approx. 5.5 m as opposed to 8.7 m). Players are free to move anywhere in the court. Yet modelling indicates that uncertainty for both players and line judges is doubled for base and service lines compared with side and centre lines. It seems unlikely that the poorer performance of both players and judges is related simply to viewing distance.

One possible factor contributing to line judge errors may be their viewing angle. A ball struck diagonally across the court from one corner to the other travels across the field of the view of a base/service line judge (sitting at the side of the court) over twice as fast as it travels across the field of view of the centre or side line judge (sitting at the end of the court, which is over twice as long as it is wide). The speed difference will be much greater for balls that are struck down the centre of the court rather than diagonally. Positional acuity deteriorates at high speed, consistent with lower performance for base/service bounces (Chung & Bedell 2003).

A second factor may be the trajectory of the ball. The contact patch of the ball with the court is elongated along the ball's trajectory. Furthermore, visible persistence in the visual system should smear the neural representation of a rapidly moving ball along its trajectory. Consequently, there should be greater uncertainty regarding bounce position parallel to the trajectory of the ball than at right angles to it. Greater uncertainty along the trajectory should have more impact on base and service line bounces, where the ball tends to travel directly across the line, than on side and centre line bounces, where the ball tends to travel more parallel to the line.

The obtained differences between the midpoints of the psychometric functions in figure 5 are not statistically reliable, and are small fractions of the estimated threshold values, so offer no support for the idea that a significant positional error contributed to judgements. It is perhaps surprising that the midpoints are so close to zero, amounting to less than one-quarter of the diameter of a tennis ball, given the laboratory studies of mis-localization mentioned earlier. Laboratory studies tend to use rather impoverished stimuli, usually very simple patterns against a uniform background. Perhaps localization errors are minimized in tennis by the use of highly practised officials who maintain fixation on the line rather than attempt to track the ball, and by the richer visual environment of a real tennis court. Indeed, performance in general is remarkably high. Line judges typically sit 5-8 m away from the nearest line on the court, and tennis balls can travel at 50 m s⁻¹ in professional matches. Yet line judges and players can judge relative ball bounce position with an accuracy of approximately 30-40 mm, a distance covered in less than 1 ms. For comparison, professional batsmen in cricket are able to locate the absolute position of a moving cricket ball with a precision of 100 mm or 5 ms (McLeod 1987; Regan 1992).

(d) Differences between players

The simple model outlined in figure 2 assumes that every call that is judged to be incorrect by the player will result in a challenge. In general, one would therefore expect higher ranked players (according to ATP world ranking at the time the tournaments were held) to make more challenges than lower ranked players because they participate in more matches. The correlation between rank (in the top 20) and total challenges per player was -0.43 (p < 0.05, d.f. = 18), broadly consistent with expectation but leaving a large amount of variance unexplained. Indeed, examination of the dataset reveals marked individual differences between players in their tendency to challenge line calls. One player in the top 20 made only seven challenges in total during all 15 tournaments, while another ranked two places below made 52 challenges. There is little indication that players making a large number of challenges were more frequently incorrect; the correlation between total challenges per player and the proportion of correct challenges was -0.06 (p=NS). On average, challenges made by the top 20 players were correct on 39% of occasions (s.d. = 11.6%), consistent with results across the whole dataset despite large variations in the actual number of challenges made by each player. It therefore seems that challenges generally reflect perceptual uncertainty rather than gamesmanship, but some players are more willing to challenge than others.

5. CONCLUSIONS

The vast majority of line-call challenges and errors in professional ATP tennis matches can be explained by a simple perceptual processing model that incorporates intrinsic positional uncertainty. Professional players and line judges are remarkably proficient at judging ball bounce position, displaying an accuracy of just a few centimetres. Ball bounces near the base and service lines are more difficult to judge than those near the side and centre lines, probably due to retinal speed of the ball and greater perceptual uncertainty along its trajectory. The model predicts that 8.2% of all line calls involving balls within 100 mm of a court line will be called incorrectly by line judges, due to inherent limitations in their perceptual system.

Some practical implications follow from these results. First, training and line judge selection should focus on maximizing performance for base and service line calls, since these are the most error prone. Second, players should attempt to make full use of all the challenges available to them because some errors are inevitable, but should bear in mind that both they and the line judges are more likely to be wrong for base and service line calls than for side and centre line calls. Finally, line-call accuracy is sufficiently high that the current rule of two unsuccessful challenges per set seems reasonable, on the following grounds. Assume that a typical set in an evenly matched professional contest involves 50 points per set (10 games with five points per game), and that each point involves a line call. Even if every call related to a ball bounce within 100 mm of a court line, the expected number of line judge errors per set is four (8.2% of calls according to the model). Hence, even in the worst-case scenario (every point ending in a borderline call), there would only be twice as many errors as there are unsuccessful challenges permitted.

I am grateful to the Association of Tennis Professionals for making their line-call challenge records available, and to Dr John Haigh of the Mathematics Department at the University of Sussex for helpful comments.

REFERENCES

- Barlow, H. B. 1956 Retinal noise and absolute threshold. J. Opt. Soc. Am. 46, 634-639.
- Bressler, D. W. & Whitney, D. 2006 Second-order motion shifts perceived position. *Vis. Res.* **46**, 1120–1128. (doi:10.1016/j.visres.2005.10.012)
- Chung, S. T. L. & Bedell, H. E. 2003 Velocity dependence of Vernier and letter acuity for band-pass filtered moving stimuli. *Vis. Res.* **43**, 669–682. (doi:10.1016/S0042-6989(02)00628-4)
- Fischetti, M. 2007 In or out? Sci. Am. July, 96-97.
- Klein, S. A. & Levi, D. M. 1987 Position sense of the peripheral retina. J. Opt. Soc. Am. A 4, 1543-1553.
- Krekelberg, B. & Lappe, M. 2000 A model for the perceived relative positions of moving objects based upon a slow averaging process. *Vis. Res.* 40, 201–215. (doi:10.1016/ S0042-6989(99)00168-6)
- Land, M. F. & Nilsson, D.-E. 2002 *Animal eyes*. Oxford, UK: Oxford University Press.
- Li, R. W., Klein, S. A. & Levi, D. M. 2006 The receptive field and internal noise for position acuity change with feature separation. *J. Vis.* 6, 311–321. (doi:10.1167/6.4.2)

- Linares, D., Lopez-Moliner, J. & Johnston, A. 2007 Motion signal and the perceived positions of moving objects. *J. Vis.* 7, 1–7. (doi:10.1167/7.7.1)
- McLeod, P. 1987 Visual reaction time and high-speed ball games. *Perception* 16, 49–59. (doi:10.1068/p160049)
- Nijhawan, R. 2002 Neural delays, visual motion and the flash-lag effect. *Trends Cogn. Sci.* **6**, 387–393. (doi:10. 1016/S1364-6613(02)01963-0)
- Oğmen, H., Patel, S. S., Bedell, H. E. & Camuz, K. 2004 Differential latencies and the dynamics of the position computation process for moving targets, assessed with the flash-lag effect. *Vis. Res.* 44, 2109–2128. (doi:10.1016/ j.visres.2004.04.003)
- Pelli, D. G. 1991 Noise in the visual system may be early. In Computational models of visual processing (eds M. Landy & J. A. Movshon), pp. 147–152. Cambridge, MA: MIT Press.

- Regan, D. 1992 Visual judgements and misjudgements in cricket, and the art offlight. *Perception* **21**, 91–115. (doi:10. 1068/p210091)
- Watt, R. J. & Hess, R. 1987 Spatial information and uncertainty in anisometropic amblyopia. Vis. Res. 27, 661–674. (doi:10.1016/0042-6989(87)90050-2)
- Westheimer, G. 1981 Visual Hyperacuity. Prog. Sensory Physiol. 1, 1-30.
- White, J. A., Rubinstein, J. T. & Kay, A. R. 2000 Channel noise in neurons. *Trends Neurosci.* 23, 131–137. (doi:10. 1016/S0166-2236(99)01521-0)
- Wichmann, F. A. & Hill, N. J. 2001 The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* 63, 1293–1313.
- Wilson, H. 1991 Model of peripheral and amblyopic hyperacuity. Vis. Res. **31**, 967–982. (doi:10.1016/0042-6989(91)90204-I)