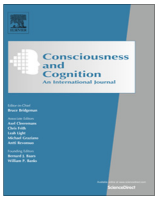
Consciousness and Cognition 27 (2014) 246–253

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog



Variance misperception explains illusions of confidence in simple perceptual decisions

Ariel Zylberberg[a,b,c,d,*], Pieter R. Roelfsema[b,e,f], Mariano Sigman[a,g]

[a] Laboratory of Integrative Neuroscience, Physics Department, FCEyN UBA and IFIBA, Conicet, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina
[b] Department of Vision and Cognition, Netherlands Institute for Neuroscience, an Institute of the Royal Netherlands Academy of Arts and Sciences, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands
[c] Institute of Biomedical Engineering, Faculty of Engineering, Buenos Aires University, 1063 Buenos Aires, Argentina
[d] Laboratory of Applied Artificial Intelligence, Computer Science Department, FCEyN UBA, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina
[e] Department of Integrative Neurophysiology, VU University Amsterdam, The Netherlands
[f] Psychiatry Department, Academic Medical Center, Amsterdam, The Netherlands
[g] Universidad Torcuato Di Tella, Almirante Juan Saenz Valiente 1010, C1428BIJ Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 24 March 2014
Available online 19 June 2014

Keywords:
Perceptual decisions
Metacognition
Signal-detection theory
Sensory reliability
Introspection

ABSTRACT

Confidence in a perceptual decision is a judgment about the quality of the sensory evidence. The quality of the evidence depends not only on its strength ('signal') but critically on its reliability ('noise'), but the separate contribution of these quantities to the formation of confidence judgments has not been investigated before in the context of perceptual decisions. We studied subjective confidence reports in a multi-element perceptual task where evidence strength and reliability could be manipulated independently. Our results reveal a confidence paradox: confidence is higher for stimuli of lower reliability that are associated with a lower accuracy. We show that the subjects' overconfidence in trials with unreliable evidence is caused by a reduced sensitivity to stimulus variability. Our results bridge between the investigation of miss-attributions of confidence in behavioral economics and the domain of simple perceptual decisions amenable to neuroscience research.

1. Introduction

Subjective confidence is used ubiquitously in approximate everyday expressions such as – "I think…", "Maybe…", "I am sure that". In signal detection theory, confidence has a precise mathematical definition, indexing the probability that the decision is actually correct (Kepecs & Mainen, 2012). A large corpus of data suggests that the brain can be close to optimal when performing perceptual inferences under uncertainty, integrating multiple sources of evidence weighted by their reliability (Knill & Pouget, 2004; Körding & Wolpert, 2006). The view that confidence is an accurate index of decision uncertainty has been a conceptual anchor for neurophysiological and psychophysical studies of confidence (Bach & Dolan, 2012). Indeed, recent studies have demonstrated how cortical circuits can represent uncertainty in perceptual decisions and how these representations guide action (Kepecs, Uchida, Zariwala, & Mainen, 2008; Kiani & Shadlen, 2009).

The emerging picture of probabilistic inference in neuroscience, however, is in striking tension with the principles of behavioral economics. Decades of experimentation with "real-life" decision problems have shown that humans are

---

[*] Corresponding author. Address: Howard Hughes Medical Institute, Department of Neuroscience and Kavli Institute for Brain Science, Columbia University New York, 630 West 168th Street, P&S Building Room 16-409, New York, NY 10032, USA. Fax: +1 212 305 9608 4510.

E-mail address: ariel.zylberberg@gmail.com (A. Zylberberg).

http://dx.doi.org/10.1016/j.concog.2014.05.012
1053-8100/

sub-optimal decision makers and, specifically, that their estimation of confidence deviates from the predictions of probability theory (Griffin & Tversky, 1992; Lichtenstein, Fischhoff, & Phillips, 1977; Slovic, 1972; Slovic & Lichtenstein, 1968; Tversky & Kahneman, 1974, 1981). In economics, humans exhibit reliable inconsistencies: they rely on sub-samples of the data, focus on tokens (representative exemplars), ignore the variance (or reliability) of the distribution, and over-weight evidence confirming their previous commitments and choices (Griffin & Tversky, 1992; Tversky & Kahneman, 1973, 1974, 1981). Consider the following example of variance neglect from behavioral economics (Kahneman, 2011). When people are presented with the sentence "In a telephone poll of 30 seniors, 60% support the president", many conclude that "most seniors support the president". This leap of confidence ignores the fact that the reliability of a poll depends on the number of participants.

The present work is an effort to reconcile theories of probabilistic inference in neuroscience with behavioral economics (Griffin & Tversky, 1992; Jarvstad, Hahn, Rushton, & Warren, 2013; Wu, Delgado, & Maloney, 2009). We investigate if subjects are able to take the variance of visual stimuli into account and if conditions exist where the perceptual choices of subjects deviate systematically from their confidence reports.

2. Results

We studied sensory decision making with a field of 60 oriented line segments. Orientations were sampled from a uniform distribution of mean orientation μ and width φ. We controlled orientation jitter across the bars (exemplars) by varying φ, which was randomly selected from a set of three values (φ_i, $i = 1, 2, 3$) (Fig. 1a). Participants reported with a single (four-choice) response whether the average orientation was right (clockwise, CW) or left (counter-clockwise, CCW) relative to the vertical axis and the confidence in their choice ('high' or 'low').

Stimuli with more orientation jitter should generate more variance in the internal representation of orientation (de Gardelle & Summerfield, 2011). We first measured the form of this intuitive relationship, because it provides insight into the mechanism by which participants estimate the mean orientation. As expected, participants' accuracy increased with signal strength ($p < 10^{-8}$, likelihood-ratio test; see methods) (Fig. 1b). In addition, the subjects' orientation sensitivity decreased with jitter φ (the interaction between $|\mu|$ and φ was significant, $p < 10^{-8}$; likelihood ratio test). To formally relate accuracy to the value of μ, we relied on signal detection theory (SDT) (Green & Sweets, 1966). Specifically, we estimated accuracy as $P_{corr}(\mu) = \int_{-\infty}^{|\mu|} N(0, \sigma)$ and determined the best-fitting internal noise level σ for each stimulus class. We obtained values of $\sigma = [0.045, 0.072, 0.11]$ radians for stimuli with low, intermediate and high orientation jitter, respectively (Fig. 1c). As expected (de Gardelle & Summerfield, 2011), the stimuli with wider distributions of orientations were associated with more internal noise.

The signal-detection theory (Green & Sweets, 1966) also provides a theoretical framework to understand how signal and noise jointly determine confidence. Confidence should depend on the ratio between signal and noise so that the criterion that separates high from low confidence trials should vary with orientation jitter (Fig. 1d). However, if subjects ignore stimulus jitter (φ) when estimating the internal noise level (σ), the criterion that separates high from low confidence trials should be fixed (Fig. 1e). For simplicity, henceforth we refer to these two alternative models as *signal–noise* and *signal* models, respectively. When cast in the framework of the telephone poll, the *signal* model predicts equal confidence in polls with 30 and 300 seniors reporting that 60% support the president, whereas the *signal–noise* model predicts higher confidence in the larger pole. Note that the *signal* model mirrors the reliability misperception (Griffin & Tversky, 1992) from behavioral economics.

To distinguish between the *signal* and the *signal–noise* model, we estimated the patterns of confidence for different values of μ and the three empirically determined values of σ_i (Fig. 1b and c). For the *signal–noise* model, a trial was considered of high confidence whenever the random sample (x) exceeded a particular threshold that increased linearly with the orientation jitter (Fig. 1d), whereas the threshold did not depend on jitter in the *signal* model (Fig. 1e). These two models generate qualitatively different predictions. In the *signal–noise* model, confidence decreases if the jitter is higher (Fig. 1f). This pattern is intuitive: subjects are less confident for stimuli sampled from broad distributions. Conversely, in the *signal* model average confidence of the subjects is actually predicted to *increase* with jitter if the signal (μ) is small (Fig. 1g). This less intuitive dependence can be understood by observing that as σ increases, the fraction of the distribution crossing this fixed criterion increases and hence the probability that a random sample is assigned high confidence increases too.

We therefore analyzed how confidence depended on the signal and the noise (Fig. 1h). The confidence of the subjects tended to be higher for trials with a larger orientation jitter. To determine significance, we measured the influence of μ and φ on confidence with a logistic regression model. As predicted by *signal* and *signal–noise* models, the proportion of high-confidence responses increased with signal strength $|\mu|$ ($p < 10^{-8}$; likelihood ratio test). The influence of $|\mu|$ on confidence was weaker if orientation jitter was higher, as can be seen in Fig. 1h where the curves become flatter for larger orientation jitter (the combined effect of $|\mu|$ and φ on confidence was evaluated including a term $|\mu| \times \varphi$ in the regression model; this interaction had a significantly negative effect on the proportion of high-confidence responses ($p < 10^{-8}$; likelihood ratio test), which indicates that higher values of φ reduce the sensitivity of confidence to changes in $|\mu|$). Jitter φ had a main effect on confidence: larger jitter leads to more confident responses ($p < 10^{-8}$). At first sight, these results support the *signal* and are incompatible with the *signal–noise* model (Fig. 1h).