The Confidence Heuristic with Bayesian Confidence

October 9, 2017

Bahrami et al. (2010) provides evidence that dyads use a confidence heuristic in 2AFC perceptive tasks, i.e. follow the choice of the most confident participant. To do so, they make subjects perform a 2AFC perceptive task in which the use of the confidence heuristic leads *unequal* dyads, i.e.dyads composed of unequally sensitive participants, to perform poorly relative to an ideal benchmark. In fact, when dyad members become extremely unequal, the confidence heuristic even leads the dyad to perform more poorly than its most sensitive member, so that the attempt to collaborate was detrimental. The observed performance of unequal dyads in Bahrami et al. (2010) matches the theoretical predictions, which suggests that they indeed used the confidence heuristic.

In the sequel, we will refer to the claim that unequal dyads using the confidence heuristic are relatively inefficient in a 2AFC perceptive task as the Inequality-Efficiency relation, or **IE relation**. Thought the IE relation is only a mean to an end in Bahrami et al. (2010), the authors suggest to study its generality.

This note points out that the IE relation as predicted by Bahrami et al. (2010) does not depend on how the experimenter generates the stimuli. We use a Bayesian model to predict the way the stimulus generating process affects the IE relation. The model suggests ways in which the IE relation could be non-monotonic or disapear altogether.

The first section briefly describes Bahrami et al. (2010)'s experiment and model. The second shows how Bahrami et al. (2010)'s results change in a Bayesian model.

1 Bahrami et al. (2010)

1.1 Perceptive Task

Participants observe two intervals, each of which contains several targets. In one of the two intervals there is an oddball target, whose contrast is slightly stronger than the other targets. The task consists in finding out which interval contains the oddball target, which is made difficult by a short exposure time. The level of contrast of the oddball target varies from a trial to the other, which makes the task harder to fulfil in some trial than in others. Each participant performing this task has a certain sensitivity, which is inferred from his success rate (the proportion of trials where the participant successfully found the interval containing the oddball target). Participants are assembled into dyads. The two dyad members observe the same stimulus, make a guess individually and then deliberate to reach a joint decision. What determines the performance of a dyad is its ability to rightly break ties, i.e. figure out which participant is more likely to be right when the two dyad members initially disagree. The confidence heuristic is an intuitive way of doing so: each participant should come up with a confidence, i.e. an assessment of her probability of being correct, and the dyad should follow its most confident member's choice.

1.2 Model

The contrast level is noted Δc and takes a positive (negative) value when the oddball target is in the right (left) interval. Δc is further away from 0 when the contrast level is stronger. The observer perceives the contrast level with some noise. Formally, he receives a signal x_i that equals Δc plus a normally distributed error, whose standard deviation σ_i reflects the observer's sensitivity. An observer's sensitivity is characterized by his sensitivity parameter s_i, which equals $1/\sigma_i$. Equipped with this signal, the observer forms a belief about where the contrast level truly stands, which is normally distributed with mean x_i and precision s_i^2. It is then used to determine the probability that the contrast level is positive or negative. If the probability that the contrast level is positive is more than half, the observer chooses the right interval and his confidence equals the probability of that Δc is positive.

Bahrami et al. (2010) presents two (main) models of the way dyad members i and j make a decision based on their signals x_i and x_j. The Direct

Signal Sharing model assumes that signals are perfectly integrated. The Weighted Confidence Sharing model assumes that dyads use the confidence heuristic, i.e. follow the choice of their most confident member.

Both models make it possible to predict a dyad's sensitivity as a function of its members' sensitivities. The sensitivity in the DSS model is

$$s_{dyad}^{DSS} = (s_i^2 + s_j^2)^{1/2}.$$

The sensitivity of the WCS model writes

$$s_{dyad}^{WCS} = \frac{s_i + s_j}{\sqrt{2}}$$

and is lower than in the DSS model whenever $s_i \neq s_j$. Following the most confident members is indeed not optimal when participants' sensitivities are not equal. The combination of individual confidences should instead be biased toward the most sensitive participant. As the difference between s_i and s_j increases, a dyad following the confidence heuristic makes more mistakes.

2 Bayesian Model

In Bahrami et al. (2010)'s model, participants' confidence does not depend on the contrast level distribution. As a consequence the experimenter could use the same contrast level all along or on the contrary generate a wide dispersion without affecting confidence formation and in turn the IE relation.

This prediction is not self-evident in an experiment featuring hundreds of trials. It is conceivable that subjects learn about the contrast level distribution from experience and use this information to form confidence. A natural way of capturing the impact of the contrast level distribution on confidence is Bayesian updating. In a Bayesian model of confidence formation, subjects hold a prior belief about the contrast level distribution that equals the actual contrast level distribution. This information is used to interpret any given signal. This note revisits the predictions of Bahrami et al. (2010) under the assumption of Bayesian updating.

2.1 Normally Distributed Contrast Levels

This section presents a Bayesian confidence model with a normally distributed prior. This model admits the model of Bahrami et al. (2010) as a special case where the prior has infinite variance.

Assume that the contrast level distribution is normal with mean 0 and precision τ. Such a distribution can be obtained in the following way: if the right interval is to contain the oddball target, the contrast level Δc is drawn from a truncated normal distribution which only takes positive values (see figure 1). Symmetrically, if the oddball target in on the left interval, the Δc is drawn from a truncated normal distribution which only takes negative values. As a result, the unconditional contrast level distribution, and in turn the prior belief, is normal with mean 0.
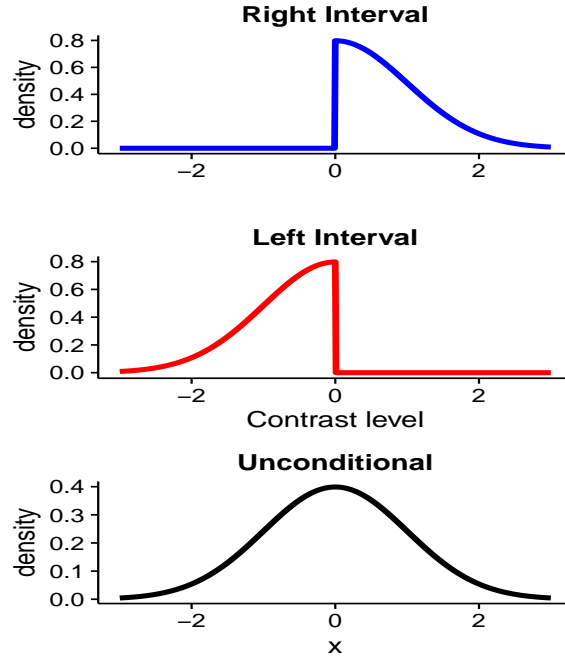

Figure 1: Distribution of Contrast Level.

Subjects update their prior belief based on their signals according to Bayes rule. The signals are normally distributed just as in Bahrami et al. (2010) so the posterior belief about Δc given x_i is normally distributed with mean

$$\frac{s_i^2}{s_i^2 + \tau} x_i$$

and precision $\tau + s_i^2$. This belief boils down to the belief in Bahrami et al. (2010) when $\tau = 0$. $\tau = 0$ corresponds to the case where the contrast level

4

distribution is infinitely variable, that is where the prior belief contains no information.

Equipped with this model of confidence formation, we derive a dyad's sensitivity under the DSS and WCS models. The details of the computations are exposed in the appendix. The DSS model gives the same sensitivity slope in our model and Bahrami et al. (2010), i.e.

$$s_{dyad}^{DSS} = (s_i^2 + s_j^2)^{1/2}.$$

Our model makes a different prediction for the WCS model however. The dyad's sensitivity indeed writes

$$d_{dyad}^{WCS} = \frac{s_i^2(\tau + s_j^2)^{1/2} + s_j^2(\tau + s_i^2)^{1/2}}{(s_i^2(\tau + s_j^2) + s_j^2(\tau + s_i^2))^{1/2}}.$$

As in Bahrami et al. (2010), the WCS model predicts a lower dyad sensitivity than the DSS model as soon as $s_i \neq s_j$. In that sense, the IE relation is still there.

The shape of this IE relation depends on τ however. Figure 2 mimicks the graph presented in Bahrami et al. (2010) that compares the WCS and DSS models. It plots s_{dyad}/s_{max} against s_{min}/s_{max} where s_min is the lowest sensitivity of the two dyad members and s_{max} the larger sensitivity of the two dyad members. The graph features the two predictions initially present in Bahrami et al. (2010), i.e. the DSS's prediction and the WCS's prediction for $\tau = 0$. The third prediction in our graph is the WCS's prediction for a positive τ, i.e. for a moderately variable stimulus strength.

Figure 2 features two interesting properties. First, it shows that the communication benefits under the WCS model become larger as τ becomes larger, i.e as the contrast level distribution becomes less variable. Second, it shows that communication benefits are not monotonic. The communication benefit curve is actually decreasing for small values of s_{min}/s_{max}.

We first give an intuition for the second prediction, which we view as less important as the first. We treat the first prediction in more depth in the next section.

The communication benefits under the confidence heuristic are not monotonic in s_{min}/s_{max}. To understand this point, note that the mistake of the confidence heuristic is to treat participants' confidences equally, regardless of participants' sensitivity. In the DSS model on the contrary, a dyad
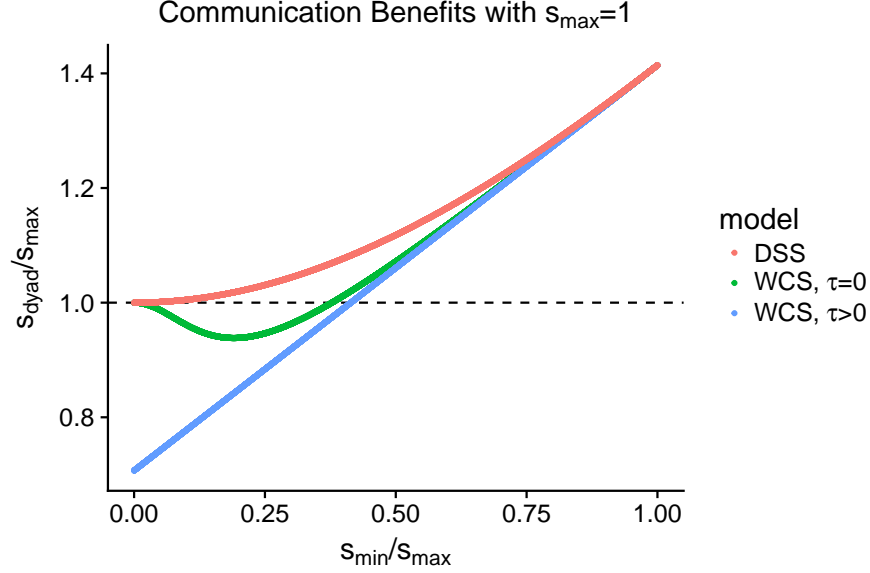
5

Figure 2: Communication Benefits

should use a threshold which is biased in the direction of the more sensitive observer and the extent of the bias increases with the extent of inequality in the dyad. Figure 3 illustrates this point. It plots the decision a dyad makes as a function of its members confidences. The area between the two curves represents all pairs of confidences for which the confidence heuristic leads the dyad to make a mistake. The figure shows that the area between the two decision curves gets larger when the difference between individual sensitivities increase. There is as a result more room for making mistakes.

On the other hand, figure 3 also reveals that the confidence heuristic makes more mistakes on intermediate confidences (around 75%) than on extreme confidences (around 50% or 100%). Since a more unequal dyad is less likely to face trials where both participants have intermediate confidences, it is also less likely to be in a situation where the confidence heuristic is prone to mistakes. Consider the extreme case where one member is not looking at the stimulus and therefore has a 50% confidence for every trial. If this dyad follows the confidence heuristic, then it will never make any mistake and its performance will equal the performance of its most sensitive participant. The reason is that this dyad is never in a situation to commit a mistake.

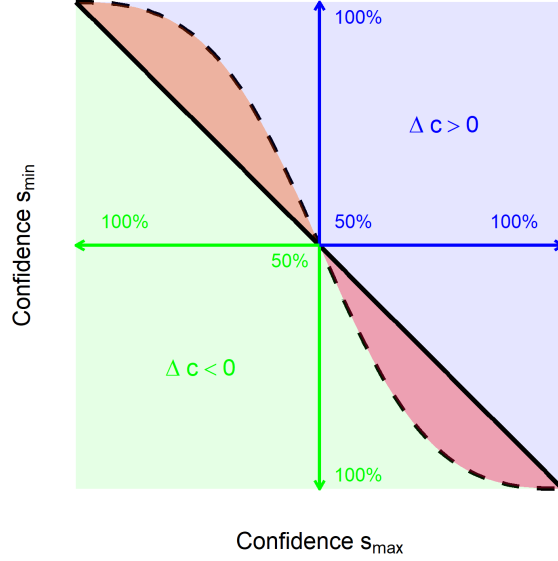The IE relation is therefore shaped by two conflicting forces. On the one

Figure 3: *Dyad's choice as a function of its members' confidence levels. The plain line represents the WCS model. The dashed line represents the DSS model. In the red area are the pairs of confidence levels for which the WCS model chooses the wrong interval.*

hand, more unequal dyads have more opportunities to make mistakes, on the other hand, they are less likely to face them. The trade-off between these two forces leads to communication benefits which are decreasing and then increasing.

2.2   Fixed Contrast Level

This section argues that in the Bayesian model of confidence formation, it is the contrast level variability that is at the root of the IE relation. Without contrast level variability, there would not be a IE relation in the first place. The model of the previous section did not allow us to clearly make this claim because the prior was normally distributed. Using a normally distributed contrast level implies that a reduction in the contrast level variability goes together with a reduced strength of the average contrast level. As a result, it mixes two effects.

Consider instead a perceptive task with a fixed contrast level. Δc only

takes two values, one positive when the oddball target in the right interval and one negative when it is in the left interval. The appropriate model for this task is the standard binormal signal detection model. As usual, individual signals x_i and x_j are drawn independently conditional on Δc. Because there is one value of Δc for each interval, the signals are also drawn independently conditional on the interval. As a consequence,

Fact 1. The confidence heuristic is the optimal rule to select an interval when the contrast level is fixed and interval are equally likely to drawn.

This fact holds regarless of the type of signal distribution, and is proven in the appendix. It follows that the IE relation disapears when the contrast level is fixed. [1]

This fact helps explain the role that the variability of the contrast level plays in the IE relation. When the contrast level is allowed to vary from a trial to another, individual signals become mechanically correlated conditional on the interval containing the oddball target. Indeed, given that the oddball target belongs to an interval, participants will both tend to receive strong signals when the contrast level is large and weak signals when the contrast level is weak. As a consequence, the confidence heuristic is not optimal anymore.

3   Conclusion

The IE relation is somewhat sensitive to the contrast level distribution when confidence is Bayesian. Whether confidence adapts to the contrast level distribution in a way consistent with Bayesian updating is not clear. So doing requires subjects to know the contrast level distribution, which can only be learned from experience.

The predictions of this note can be tested in the lab. This can be done by running experiments with dyadic interaction and different contrast level distribution. Another way is to collect subjects' confidence and form hypothetical dyads on which dyad choices are made following the confidence heuristic.

---

[1]This result assumes that individual signals are independently drawn conditional on a contrast level. This assumption is rather standard but might hold in some perceptive tasks.

References

B. Bahrami, K. Olsen, P. E. Latham, A. Roestorff, G. Rees, C. D. Frith. "Optimally Interacting Minds." *Science, Aug. 2010 : 1081-1085.*

Appendix

This section presents the formal derivations of some of the note's results. In the sequel, whenever we write that a random variable follows a $\mathcal{N}(\cdot, \cdot)$, the second parameter will be the precision, i.e. the inverse of the variance.

Derivation of d_{dyad}^{WCS}

The prior distribution of Δc is $\mathcal{N}(0, \tau)$. Given that signal x_i's distribution given Δc is $\mathcal{N}(\Delta c, s_i^2)$, the posterior belief about Δc given x_i is

$$\mathcal{N}(\frac{s_i^2}{s_i^2 + \tau}, \tau + s_i^2).$$

It follows that the posterior probability that $\Delta c \geq 0$ given x_i equals

$$\Phi(\frac{s_i^2}{\sqrt{s_i^2 + \tau}} x_i)$$

where Φ is the standard normal cdf.

If a dyad composed of i and j follows the most confident member, it chooses the right interval if the posterior probability of $\Delta c \geq 0$ given x_i is larger than the posterior probability of $\Delta c \leq 0$ given x_j. The decision rule writes

$$\frac{s_i^2}{\sqrt{\tau + s_i^2}} x_i + \frac{s_j^2}{\sqrt{\tau + s_j^2}} x_j \geq 0.$$

Given a contrast level Δc, this decision statistic is normally distributed with mean

$$(\frac{s_i^2}{\sqrt{\tau + s_i^2}} + \frac{s_j^2}{\sqrt{\tau + s_j^2}}) \Delta c$$

and variance

$$\frac{s_i^2}{\tau + s_i^2} + \frac{s_j^2}{\tau + s_j^2}.$$

9

So, the dyad's sensitivity writes

$$d_{dyad}^{WCS} = \frac{s_i^2(\tau + s_j^2)^{1/2} + s_j^2(\tau + s_i^2)^{1/2}}{(s_i^2(\tau + s_j^2) + s_j^2(\tau + s_i^2))^{1/2}}.$$

Derivation of d_{dyad}^{DSS}

The prior distribution of Δc is $\mathcal{N}(0, \tau)$. Given that x_i and x_j are both normally distributed with mean Δc and precisions s_i^2 and s_j^2 respectively, the dyad's posterior belief about Δc is normally distributed with mean

$$\frac{s_i^2 x_i + s_j^2 x_j}{\tau + s_i^2 + s_j^2}$$

and precision $\tau + s_i^2 + s_j^2$.

The dyad then chooses the right interval if its posterior mean is positive, that is

$$s_i^2 x_i + s_j^2 x_j \geq 0.$$

This decision statistic is normally distributed with mean $\Delta c(s_i^2 + s_j^2)$ and variance $s_i^2 + s_j^2$. The dyad's sensitivity then writes

$$s_{dyad}^{DSS} = (s_i^2 + s_j^2)^{1/2}.$$

The confidence heuristic is optimal with a fixed contrast level and two equally likely intervals.

The contrast level Δc takes a value θ with probability $\pi(\theta)$ and $-\theta$ with probability $\pi(-\theta)$.

Participants i and j receive independent signals conditional on Δc. The density functions of x_i and x_j are noted $f_i(\cdot|\Delta c)$ and $f_j(\cdot|\Delta c)$ respectively. The joint density of x_i and x_j is noted $f(\cdot|\Delta c)$, which equals $f_i(\cdot|\Delta c)f_j(\cdot|\Delta c)$.

If the dyad uses the confidence heuristic, it chooses $\Delta = \theta$ if

$$Pr(\theta|x_i) \geq Pr(-\theta|x_j).$$

This inequality can be rewritten as

$$\frac{\pi(\theta)f_i(x_i|\theta) + \pi(-\theta)f_i(x_i|-\theta)}{f_i(x_i|\theta)\pi(\theta)} \leq \frac{\pi(\theta)f_j(x_j|\theta) + \pi(-\theta)f_j(x_j|-\theta)}{f_j(x_j|\theta)\pi(-\theta)}$$

that is

$$\frac{f_i(x_i|-\theta)\pi(-\theta)}{f_i(x_i|\theta)\pi(\theta)} \leq \frac{f_j(x_j|\theta)\pi(\theta)}{f_j(x_j|\theta)\pi(-\theta)}.$$

If the dyad optimally integrates its signals, then it chooses θ if

$$Pr(\theta|x_i, x_j) \geq Pr(-\theta|x_i, x_j).$$

Provided that signals are independent conditional on Δc, this inequality can be rewritten as

$$\frac{f_i(x_i|\theta)f_j(x_j|\theta)\pi(\theta)}{f(x_i, x_j)} \geq \frac{f_i(x_i|-\theta)f_j(x_j|-\theta)\pi(-\theta)}{f(x_i, x_j)},$$

which is equivalent to

$$\frac{f_j(x_j|\theta)\pi(\theta)}{f_j(x_j|-\theta)} \geq \frac{f_i(x_i|-\theta)\pi(-\theta)}{f_i(x_i|\theta)}.$$

The two decision rules are equivalent as long as $\pi(\theta) = \pi(-\theta)$. [2]

_____

[2]So the Bayesian model actually predicts that if an interval is chosen more often than the other one, the confidence heuristic will be inefficient.