



# Context Effects in the Judgment of Visual Relative-Frequency: Trial-by-Trial Adaptation and Non-linear Sequential Effect

Xiangjuan Ren<sup>1,2</sup>, Muzhi Wang<sup>3</sup> and Hang Zhang<sup>2,3,4\*</sup>

<sup>1</sup> Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China, <sup>2</sup> Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China, <sup>3</sup> School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China, <sup>4</sup> PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

## OPEN ACCESS

### Edited by:

Andrey R. Nikolaev,  
KU Leuven, Belgium

### Reviewed by:

Guido Marco Cicchini,  
Consiglio Nazionale Delle Ricerche  
(CNR), Italy  
Ambarish Pawar,  
Salk Institute for Biological Studies,  
United States  
Ingo Fründ,  
York University, Canada

### \*Correspondence:

Hang Zhang  
hang.zhang@pku.edu.cn

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

Received: 10 April 2018

Accepted: 22 August 2018

Published: 12 September 2018

### Citation:

Ren X, Wang M and Zhang H (2018)  
Context Effects in the Judgment of  
Visual Relative-Frequency:  
Trial-by-Trial Adaptation and  
Non-linear Sequential Effect.  
Front. Psychol. 9:1691.  
doi: 10.3389/fpsyg.2018.01691

Humans' judgment of relative-frequency, similar to their use of probability in decision-making, is often distorted as an inverted-S-shape curve—small relative-frequency overestimated and large relative-frequency underestimated. Here we investigated how the judgment of relative-frequency, despite its natural reference points (0 and 1) and stereotyped distortion, may adapt to the environmental statistics. The task was to report the relative-frequency of black (or white) dots in a visual array of black and white dots. We found that participants' judgment was distorted in the typical inverted-S-shape, but the distortion curve was influenced by both the central tendency and spread of the distribution of objective relative-frequencies: the lower the central tendency, the higher the overall judgment (contrast effect); the higher the spread, the more curved the inverted-S-shape (curvature effect). These context effects are in the spirit of efficient coding but opposite to what would be predicted by Bayesian inference. We further modeled the context effects on the level of individual trials, through which we found not only a trial-by-trial adaptation, but also the non-linear sequential effects that were recently reported mainly in circularly distributed visual stimuli.

**Keywords:** probability distortion, subjective probability, frequency estimation, sequential effect, adaptation, Bayesian inference, efficient coding

## INTRODUCTION

Humans' judgment of relative-frequency, similar to their use of probability in decision-making, is often distorted as an inverted-S-shape curve—small relative-frequency overestimated and large relative-frequency underestimated. Here we investigated how the judgment of relative-frequency, despite its natural reference points (0 and 1) and stereotyped distortion, may adapt to the environmental statistics. The task was to report the relative-frequency of black (or white) dots in a visual array of black and white dots. We found that participants' judgment was distorted in the typical inverted-S-shape, but the distortion curve was influenced by both the central tendency and spread of the distribution of objective relative-frequencies: the lower the central tendency, the higher the overall judgment (contrast effect); the higher the spread, the more curved the inverted-S-shape (curvature effect). These context effects are in the spirit of efficient coding but opposite to what would be predicted by Bayesian inference. We further modeled the context effects on the level of individual trials, through which we found not only a trial-by-trial adaptation, but also the non-linear sequential effects that were recently reported mainly in circularly distributed visual stimuli.

$\lambda[\pi(p)] = \gamma\lambda[p] + (\mathbb{C} - \gamma)\lambda[p]$

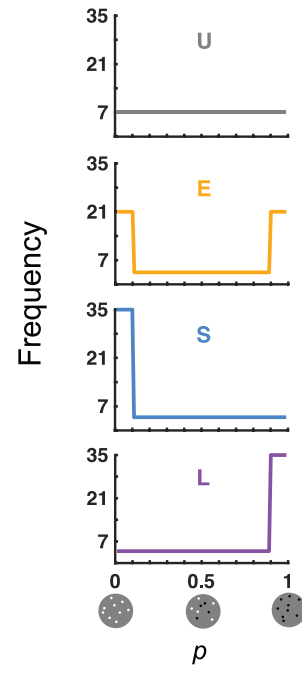
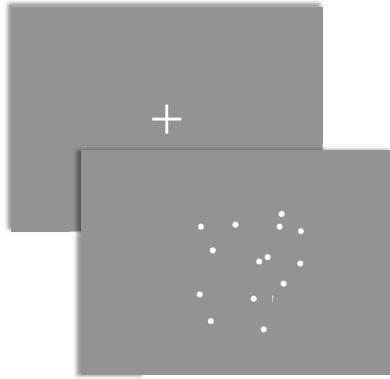
Figure 1A  
 Figure 1B

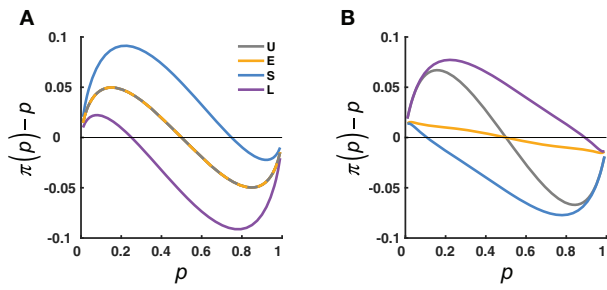
Figure 2  
 Figure 2A  
 Figure 2B

$\lambda[\pi(p)] = \gamma\lambda[p] + (\mathbb{C} - \gamma)\lambda[p]$

Figure 1A  
 Figure 1B

Figure 2  
 Figure 2A  
 Figure 2B





**FIGURE 2** | Opposing effects predicted by two influential lines of theories. The predicted deviation of the subjective from objective relative-frequency,  $\pi(\rho) - \rho$ , is plotted as a function of the objective relative-frequency,  $\rho$ , and compared across the four distribution conditions (U, E, S, L, color coded). **(A)** Adaptation-level theory (Helson, 1947). The  $\pi(\rho)$  is assumed to reflect the

For Figure 2B  $\sigma_{noise} = \mathbb{K}$

$$\pi_y = \sum_q q \mathbb{I}_r(q|y)$$

For Figure 2A  $\sigma_{noise} = \mathbb{K}$

$$\pi(p) = \int \pi_y \mathbb{I}_r(y|p) dy$$

For Figure 2B  $\sigma_{noise} = \mathbb{K}$

### Adaptation-Level Theory

For Figure 2A  $\sigma_{noise} = \mathbb{K}$

$$\lambda[\pi(p)] = \gamma(\lambda[p] - \lambda.L) + (\mathbb{K} - \gamma)\lambda[p]$$

For Figure 2A  $\sigma_{noise} = \mathbb{K}$

$$\lambda.L = \eta \sum_p \theta(p) \lambda[p]$$

For Figure 2A  $\sigma_{noise} = \mathbb{K}$

$$\text{For Figure 2A } \sigma_{noise} = \mathbb{K} \quad \gamma = \quad p = \quad \eta =$$

## Measures of Distortion of Relative-Frequency

### Slope and Crossover Point Estimated From LLO

For Figure 2A  $\sigma_{noise} = \mathbb{K}$

### Smoothed Distortion Curve and Non-parametric Measures

For Figure 2A  $\sigma_{noise} = \mathbb{K}$

$$\hat{M}_h(x) = \frac{\sum_{i=1}^m K\left(\frac{x-x_i}{h}\right) y_i}{\sum_{i=1}^m K\left(\frac{x-x_i}{h}\right)}$$

For Figure 2A  $\sigma_{noise} = \mathbb{K}$

Let  $\sigma_k$  denote the probability of a response of  $k$  on trial  $n$ . Let  $p_j$  denote the probability of a response of  $j$  on trial  $n$ . Let  $R_n$  denote the response on trial  $n$ . Let  $S_n$  denote the number of responses of  $k$  on trial  $n$ . Let  $R_N$  denote the number of responses of  $k$  on trial  $N$ . Let  $Y$  denote the vector of responses on trial  $n$ . Let  $X$  denote the matrix of responses on trial  $n$ . Let  $W$  denote the matrix of weights on trial  $n$ .

$$W = \begin{pmatrix} w_1(p_n, p_{n-1}) & \dots \\ w_2(p_n, p_{n-1}) & \dots \\ \dots & \dots \\ w_N(p_n, p_{n-1}) & \dots \end{pmatrix}$$

$$X = \begin{pmatrix} S_1 & R_{1k} \\ S_2 & R_{2k} \\ \dots & \dots \\ S_N & R_{Nk} \end{pmatrix}$$

$$Y = \begin{pmatrix} R_1 \\ R_2 \\ \dots \\ R_N \end{pmatrix}$$

The weighted least squares (WLS) estimator of the parameters  $\beta$  is given by

$$\begin{pmatrix} \hat{\beta}^{WLS} \\ \hat{\beta}_C^{WLS} \end{pmatrix} = (X^T W X)^{-1} X^T W Y$$

### Modeling

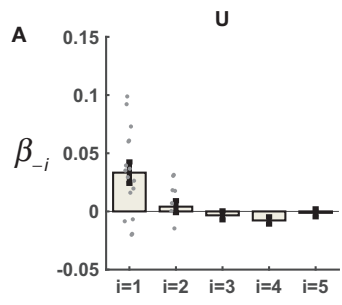
Let  $\pi(p)$  denote the probability of a response of  $k$  on trial  $n$  given the probability of a response of  $j$  on trial  $n-1$ . Let  $R_n$  denote the response on trial  $n$ . Let  $\lambda$  denote the parameter of the model. Let  $\pi(p_n)$  denote the probability of a response of  $k$  on trial  $n$  given the probability of a response of  $j$  on trial  $n-1$ .

TABLE 1 | Notations.

VARIABLES OR FUNCTIONS	
$\rho$	A generic value on the probability scale; objective probability or relative-frequency
$\lambda(\cdot)$	Log-odds function of probability or relative-frequency. $\lambda(\rho) = \log(\rho/(1-\rho))$
$\pi(\rho)$	Subjective probability or relative-frequency
$\rho_n$	Objective probability or relative-frequency of Trial $n$
$S_n$	Stimulus of Trial $n$ in log-odds. $S_n = \lambda(\rho_n) = \log(\rho_n/(1-\rho_n))$
$R_n$	Response of Trial $n$ in log-odds. $R_n = \lambda(\pi[\rho_n]) = \log(\pi[\rho_n]/(1-\pi[\rho_n]))$
MODEL ABBREVIATIONS	
LLO	Linear in log-odds model
AL	Adaptation-level model
LLO-L	Linear in log-odds model with linear sequential effects
LLO-NL	Linear in log-odds model with non-linear sequential effects
AL-L	Adaptation-level model with linear sequential effects
AL-NL	Adaptation-level model with non-linear sequential effects
MODEL PARAMETERS	
$\gamma$	Slope of the linear transformation of log-odds
$\rho_0$	Crossover point; controlling the intercept of the linear transformation of log-odds
$\beta_0$	Coefficient for the $S_n$ term
$\beta_{-i}$	Coefficient for the $R_{n-i}$ term, $i = 1, 2, \dots, 5$
$\beta_C$	Coefficient for the constant term
$\sigma_{noise}$	Standard deviation of the Gaussian noise







**TABLE 2** | Pearson's  $r$  between data and model predictions in the pattern of sequential effects.

Condition	LLO	AL	LLO-L	AL-L	LLO-NL	AL-NL
Uniform	-0.015	-0.048	-0.215	-0.217	0.773	0.668
Extreme	-0.355	-0.174	-0.126	-0.318	0.723	0.588
Small	-0.141	0.219	-0.367	0.120	0.709	0.666
Large	0.238	-0.046	0.137	0.041	0.577	0.657

### Model Comparison

o co p n t or r n e n n r o p r t r  
t n o co p t A n or t on cr r on  
corr t or p A C e A r e  
n

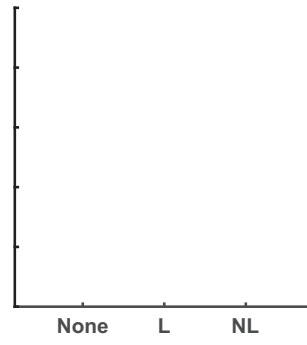
$$AIC_c = -n \ln(\hat{L}) + k + \frac{k(k-1)}{N-k-1}$$

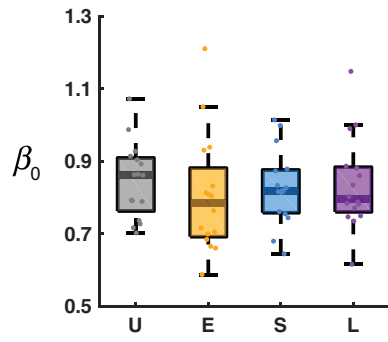
or e p p n e o t e or oo n  
o r n  $(\hat{L})$  no t oo k  
no t n r o p r t r N no t n r  
o t o r A C e t r o t  
o on t o A o  
or t on con on e p or E l con on  
r t on n con t A  
ecor n A C e l ero p p h . Figure 5  
A ro p B n o on t p n t  
on A C e t t  
r o n Figure 5 or pro t e ne pro t  
t t pro t p e o t r n t  
d r o e n o t o non  
n r q n t o t r or o n r or non  
q n d r t n n t E e p or t  
E t con on A o t r n o  
n o A o r o n n or  
con on n n t E e con on pro  
e t on o r r q ne e r r t  
p = n con on r t n p on  
r t ro t n t

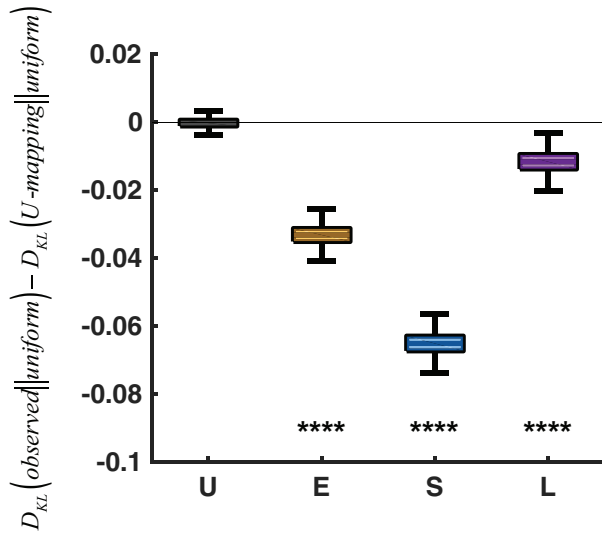
### Estimated Parameters

t t p r t r or t A o r o n n  
Figure 6 p r t r e n n o r e t or  
e on t o n Table 1 or or t t op n  
n r e p p r t r t on t or n o  $\beta$   
 $\beta_C$  r n n r t o p on  $\kappa$  t p r t r  
t con o non n r q n |  $\beta$   $\omega$  n t  
t n r t | | | | | |

Summed  $\Delta AICc$







**FIGURE 7** | Evidence for efficient coding.  $D_{KL}(\text{observed} \parallel \text{uniform})$  denotes the KL divergence from the uniform distribution over  $[0, 1]$  to the observed response distribution of a specific condition.  $D_{KL}(\text{U-mapping} \parallel \text{uniform})$  denotes the KL divergence from the uniform distribution to the response distribution predicted by the  $\rho$ -to- $\pi$  ( $\rho$ ) mapping of the Uniform condition.  $D_{KL}(\text{observed} \parallel \text{uniform}) - D_{KL}(\text{U-mapping} \parallel \text{uniform})$  is plotted for each condition, with the Uniform condition serving as a sanity check (i.e., the difference should be 0). A negative difference implies that the  $\rho$ -to- $\pi$  ( $\rho$ ) mapping adopted by a specific condition is closer to efficient coding than applying the mapping of the Uniform condition to the condition. In the box plot, the middle line denotes the median across 1,000,000 simulations, the bottom and top lines denote the lower and upper quartiles, and the error bars denote the 99% confidence interval. \*\*\*\*

Bo... n A... n A... n A... n A... n D... n... n... on or n r ne...  
n *Applied Smoothing Techniques for Data Analysis: The Kernel Approach  
With S-plus Illustrations*, ... A C A... n on B Cop D A ... re...  
... e... n D... n... n... n... n... n... n... n... n... n... n... n...  
or ...  
Br n r D... p e op e... o *Spat. Vis.* ...  
o ... X  
Broo... B... n... A... Error p... n... n... n... n... n...  
n pro... on o n... r e... propo... on *Percept. Psychophys.* ...  
o ... .BF ...  
B rr D n C e e n... on... n... p... e o n... *Curr. Biol.*  
... o ...  
B rr D n... A... n... n... o n... r *Curr. Biol.* ...  
o ...  
C op n A... n... n... n... n... n... n... n... n... n... n... n... n... n... n... n...  
*Curr. Biol.* ... o ...  
C e e n... An o... n... B rr D C... Co pr... pp n...  
o n... r... p... e... r... n... e... neo n... e... n... n... n... n... n... n... n...  
o... n... e... n... or... *Proc. Natl. Acad. Sci. U.S.A.* ...  
o... .pn ...  
C e e n... Arr... C... e... n... B rr D C...  
p... neo n... n... n... n... n... n... n... n... n... n... n... n... n... n... n... n...  
... o... . E... G...  
Co n D... Co... e... n... on n h... ro p

X X n t e r A A . E t i c o n p r o r t n  
 n p r o r n o o n p r e p B n n r n e n N I P S ' 1 2  
 P r o c e e d i n g s o f t h e 2 5 t h I n t e r n a t i o n a l C o n f e r e n c e o n N e u r a l I n f o r m a t i o n  
 P r o c e s s i n g S y s t e m s .  
 X X n t e r A A . A B n o r r o c o n t n  
 t i c o n e n p n h B n p r e p N a t . N e u r o s c i .  
 o n D B r o n  
 c o p o n n p r e p P L o S C o m p u t . B i o l .  
 o o r n p  
 D o n o n E c o n o m i c  
 e o n n e c o p r n q r o r Proc.  
 N a t l . A c a d . S c i . U . S . A .  
 o o . p n

Z n n o n q o o c o o n  
 r p r o n o p r o t n r q n e o n p r e p o n  
 n c o n F r o n t . N e u r o s c i . o . n n

**Conflict of Interest Statement:** or e r t r e  
 c o n n n e o n c o r e o r n n e r o n p t c o  
 c o n p o n c o n t o r t

Copyright © 2018 Ren, Wang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.